

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:38:35 ; Search time 28.8525 Seconds
(without alignments)
92.367 Million cell updates/sec

Title: US-09-931-325C-3
Perfect score: 115
Sequence: 1 NANPNVDPNANPNANPNANP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A.Geneseq_101002.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	115	100.0	20	23	AAU93823	P. falciparum CS B
2	115	100.0	20	23	AAU87697	Malaria B cell ep
3	115	100.0	23	23	AAU93868	P. falciparum CS-r
4	115	100.0	23	23	AAU87723	Malaria B cell ep
5	115	100.0	24	23	AAU03175	Synthetic NANP pep
6	115	100.0	24	23	AAU93822	P. falciparum CS B
7	115	100.0	24	23	AAU87696	Malaria B cell ep
8	115	100.0	28	23	AAU93825	P. falciparum CS B
9	115	100.0	28	23	AAU87699	Malaria B cell ep
10	115	100.0	31	23	AAU93867	P. falciparum CS-r

11	115	100.0	31	23	AAU93870	P. falciparum CS-r
12	115	100.0	31	23	AAU87722	Malaria B cell ep
13	115	100.0	31	23	AAU87725	Anti-Malaria Pept
14	115	100.0	38	9	AAU82566	P. falciparum cirtc
15	115	100.0	126	19	AAU9273	Antigenic determin
16	115	100.0	131	21	AAU80465	R32et32 CS sequen
17	115	100.0	133	9	AAU80466	Antigenic determin
18	115	100.0	136	8	AAU70069	Sequence of conj
19	115	100.0	136	8	AAU70069	Chimeric Hepatitis
20	115	100.0	195	23	AAU93875	Malaria antigen R
21	115	100.0	212	21	AAU15662	Synthetic peptide
22	115	100.0	250	7	AAU60412	Sequence of conj
23	115	100.0	264	8	AAU70070	Sequence of conj
24	115	100.0	267	9	AAU70071	Sequence encoded b
25	115	100.0	411	9	AAU83144	CS protein of mala
26	115	100.0	412	7	AAU83146	Sequence encoded b
27	115	100.0	412	9	AAU80835	Sequence which def
28	110	95.7	133	9	AAU80392	Malaria parasite P
29	106	92.2	20	19	AAU69278	P. falciparum cirtc
30	106	92.2	46	19	AAU59274	Immunogenic branch
31	106	92.2	56	22	AAU63514	Plasmodium falcipa
32	106	92.2	59	22	AAU63561	Sequence of a port
33	106	92.2	59	22	AAU63513	Recombinant protei
34	106	92.2	63	12	AAU14263	Sequence of conj
35	106	92.2	64	12	AAU14261	Immunogenic branch
36	106	92.2	65	12	AAU14265	Immunogenic branch
37	106	92.2	65	12	AAU14262	Immunogenic branch
38	106	92.2	66	16	AAU6442	Plasmodium falcipa
39	106	92.2	67	7	AAU60900	Sequence of a port
40	106	92.2	68	19	AAU44252	Recombinant protei
41	106	92.2	68	19	AAU44252	Sequence of conj
42	106	92.2	70	8	AAU70066	Sequence of conj
43	106	92.2	70	8	AAU70067	Sequence of conj
44	106	92.2	70	8	AAU70068	Sequence of conj
45	106	92.2	72	7	AAU60942	Sequence of a port

ALIGNMENTS

RESULT 1
AAU93823 standard; Peptide: 20 AA.
ID AAU93823:
AC AAU93823:
XX 02-JUL-2002 (first entry)
DT XX
XX P. falciparum CS B cell epitope #3.
DE XX
XX Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; HBC;
KW vaccine; B cell epitope; T cell epitope; immunostimulant.
XX Plasmodium falciparum.
OS XX
XX WO200214478-A2.
PN XX
PD 21-FEB-2002.
XX 16-AUG-2001; 2001WO-US41759.
XX 16-AUG-2000; 2000US-225843P.
PR 22-AUG-2000; 2000US-226867P.
PR 15-AUG-2001; 2001US-0930915.
PA (APOV-) APOVIA INC.
XX Birkett AJ.
XX WPI; 2002-257601/30.
DR Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric
XX hepatitis B core protein, displays immunogenic epitopes at N-terminus,
PT

PT Hbc immunogenic loop with linker for conjugated epitope and C-terminus
PT
XX
PS Example 4; Page 121; 289pp; English.
XX
CC The invention relates to a recombinant hepadnavirus nucleocapsid protein,
CC i.e. a chimeric hepatitis B core (Hbc) protein (1), displaying one or
CC more immunogenic epitopes at the N-terminus, Hbc immunogenic loop (L) or
CC C-terminus, or having a heterologous linker for a conjugated epitope in
CC (L), and containing a Cys residue at, or near, the C-terminus that
CC confers enhanced stability to the particles. A vaccine comprising (1) is
CC useful for inducing an immune response in an inoculated host animal, by
CC inoculating a host animal with the vaccine, and maintaining that
CC inoculated animal for a time period sufficient for that animal to
CC develop an immune response. The immunogenic particles formed using (1)
CC are substantially free of binding to nucleic acids, and are most stable
CC than the particle formed from otherwise identical Hbc chimera that lacks
CC the C-terminal residue or in which a C-terminal Cys is replaced by
CC another residue. The chimera particles are most stable on storage in
CC aqueous compositions that are particles of similar sequence that lack any
CC C-terminal Cys residues. The chimera molecule exhibits the self-assembly
CC not exhibiting the nucleic acid binding of those native particles, and
CC excellent B cell and T cell immunogenicities. The chimera particles are
CC typically prepared in higher yield than similar particles that are free
CC of a C-terminal Cys. The particles are often far more immunogenic than
CC the similar conjugates that lack a C-terminal Cys. Immunogenicities of
CC particles assembled from the chimera molecules are enhanced as compared to
CC similar particles assembled from chimera molecules lacking at least one
CC C-terminal Cys. AAU93802-AAU93997 represent immunogenic Hbc particles
CC amino acid sequences and related sequences of the invention.
XX
SQ Sequence 20 AA:
Query Match 100.0%; Score 115; DB 23; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NANPNVDPNANPNANPNANP 20
DQ 1 NANPNVDPNANPNANPNANP 20
RESULT 2
AAU87697
ID AAU87697 standard; Peptide: 20 AA.
XX
AC AAU87697;
XX
XX 21-MAY-2002 (first entry)
DE Malarial B cell epitope #3.
XX
XX Hepatitis B virus; nucleocapsid protein; Hbc; hepatitis B virus core;
KM B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoa;
KM Circumsporozoite; human immunodeficiency virus type 1; human; squirrel;
KM woodchuck.
XX
XX Plasmodium falciparum.
OS
PN WO200213765-A2.
PD 21-FEB-2002.
XX
XX 16-AUG-2001; 2001WO-US25625.
PF
XX 16-AUG-2000; 2000US-225813P.
PR 15-AUG-2001; 2001US-0931325.
XX
XX (APOV-) APOVIA INC.
PA
XX Birkett AJ;
PI
XX WPI; 2002-241832/29.

XX
PT Recombinant hepatitis B virus core protein chimera molecule, useful to
PT induce antibodies to malarial parasites, comprises malaria-specific
PT T-cell epitope and is engineered for enhanced stability -
XX
PS Claim 49; Page 31; 197pp; English.
XX
XX The invention relates to a recombinant hepatitis B virus core (Hbc)
CC protein chimera molecule that contains 4 peptide-linked amino acid residue
CC sequence domains. The molecule of the invention contains a region
CC constituting a B cell epitope of the circumsporozoite protein of a
CC species of the parasite, Plasmodium. The chimera sequence is useful as an
CC immunogen for inducing antibodies to the malaria-causing parasite,
CC Plasmodium, particularly P. falciparum and P. vivax. Sequences
CC AAU87695-AAU87804 represent peptide epitopes of the invention.
XX
SQ Sequence 20 AA:
Query Match 100.0%; Score 115; DB 23; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NANPNVDPNANPNANPNANP 20
DQ 1 NANPNVDPNANPNANPNANP 20
RESULT 3
AAU93868
ID AAU93868 standard; Peptide: 23 AA.
XX
AC AAU93868;
XX
XX 02-JUL-2002 (first entry)
DT
XX P. falciparum CS-repeat B cell epitope Pf3.1.
DE
XX Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; Hbc;
KM vaccine; B cell epitope; T cell epitope; immunostimulant.
XX
XX Plasmodium falciparum.
OS
XX WO200214478-A2.
PN
XX 21-FEB-2002.
PD
XX 16-AUG-2001; 2001WO-US41759.
PF
XX 16-AUG-2000; 2000US-225843P.
PR 22-AUG-2000; 2000US-226867P.
PR 15-AUG-2001; 2001US-0930915.
XX
XX (APOV-) APOVIA INC.
PA
XX Birkett AJ;
PI
XX WPI; 2002-257601/30.
DR
XX Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric
PT hepatitis B core protein, displays immunogenic epitopes at N-terminus,
PT Hbc immunogenic loop with linker for conjugated epitope and C-terminus
PT
XX
PS Example 1; Page 106; 289pp; English.
XX
XX The invention relates to a recombinant hepadnavirus nucleocapsid protein,
CC i.e. a chimeric hepatitis B core (Hbc) protein (1), displaying one or
CC more immunogenic epitopes at the N-terminus, Hbc immunogenic loop (L) or
CC C-terminus, or having a heterologous linker for a conjugated epitope in
CC (L), and containing a Cys residue at, or near, the C-terminus that
CC confers enhanced stability to the particles. A vaccine comprising (1) is
CC useful for inducing an immune response in an inoculated host animal, by
CC inoculating a host animal with the vaccine, and maintaining that

CC inoculated animal for a time period sufficient for that animal to
 CC develop an immune response. The immunogenic particles formed using (I)
 CC are substantially free of binding to nucleic acids, and are most stable
 CC than the particle formed from otherwise identical Hbc chimera that lacks
 CC the C-terminal residue or in which a C-terminal Cys is replaced by
 CC another residue. The chimera particles are most stable on storage in
 CC aqueous compositions that are particles of similar sequence that lack any
 CC C-terminal Cys residues. The chimera molecule exhibits the self-assembly
 CC not exhibiting the nucleic acid binding of those native particles, and
 CC excellent B cell and T cell immunogenicities. The chimera particles are
 CC typically prepared in higher yield than similar particles that are free
 CC of a C-terminal Cys. The particles are often far more immunogenic than
 CC the similar conjugates that lack a C-terminal Cys. Immunogenicities of
 CC particles assembled from the chimera molecules are enhanced as compared to
 CC similar particles assembled from chimera molecules lacking at least one
 CC C-terminal Cys. AAU93802-AAU93997 represent immunogenic Hbc particles
 CC amino acid sequences and related sequences of the invention.

SQ Sequence 23 AA;

Query Match 100.0%; Score 115; DB 23; Length 23;

Best Local Similarity 100.0%; Pred. No. 3.7e-08; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNVDPNANPNANPNANP 20

DB 2 NANPNVDPNANPNANPNANP 21

RESULT 4

AAU87723

ID AAU87723 standard; Peptide; 23 AA.

AAU87723;

21-MAY-2002 (First entry)

DE Malarial B cell epitope #26.

XX Hepatitis B virus; nucleocapsid protein; Hbc; hepatitis B virus core;

KW B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoa;

KM circumsporozoite; human immunodeficiency virus type I; human; squirrel;

XX woodchuck.

OS Plasmodium falciparum.

PN WO200213765-A2.

21-FEB-2002.

16-AUG-2001; 2001MO-US25625.

16-AUG-2000; 2000US-225813P.

15-AUG-2001; 2001US-0931325.

(APOV-) APOVIA INC.

Birkett AJ;

WPI; 2002-241832/29.

N-PDB; ABK44224.

Recombinant hepatitis B virus core protein chimera molecule, useful to

induce antibodies to malarial parasites, comprises malaria-specific

T-cell epitope and is engineered for enhanced stability -

CC immunogen for inducing antibodies to the malaria-causing parasite,

CC Plasmodium, particularly P. falciparum and P. vivax. Sequences

CC AAU87695-AAU87804 represent peptide epitopes of the invention.

XX SQ Sequence 23 AA;

OY 1 NANPNVDPNANPNANPNANP 20

DB 2 NANPNVDPNANPNANPNANP 21

RESULT 5

AAU03175

ID AAU03175 standard; Peptide; 24 AA.

AAU03175;

24-OCT-2001 (First entry)

Synthetic NANP peptide coupled to filamentous bacteriophage.

XX Filamentous bacteriophage FI; immunogenic carrier; immunisation;

KW anti-peptide antibody generation; major coat protein pVIII; NANP.

XX Synthetic.

XX Key

XX Region

XX Modified-size

XX /label= Orn

XX /note= "Biotinylated ornithine"

XX CA2286301-A1.

XX 01-MAY-2001.

XX 01-NOV-1999; 99CA-2286301.

XX 01-NOV-1999; 99CA-2286301.

XX (ZMIC/) ZWICK M B.

XX (SCOT/) SCOTT J K.

XX ZWICK MB, Scott JK;

XX WPI; 2001-375134/40.

XX Use of the filamentous bacteriophage as a common immunogenic carrier

XX for generating anti-peptide antibodies against recombinant peptides

XX chemically coupled to its surface, useful as a vaccine against the

XX peptides -

XX Disclosure; Page 13; 42pp; English.

XX The present sequence for synthetic NANP peptide is coupled

XX to the filamentous bacteriophage FI. The sequence is described in

XX an invention relating to the novel use of the filamentous

XX bacteriophage as a common immunogenic carrier for generating

XX anti-peptide antibodies against recombinant peptides displayed on

XX its surface. The recombinant peptides are chemically coupled to the

XX phage surface. The filamentous bacteriophages can be used as a vaccine

XX against the peptides displayed on its surface. For immunisations,

XX peptides are displayed as fusions to the major coat protein pVIII. The

XX amount of peptide that is covalently linked to the phage is increased

XX 2-fold when a lysine residue is engineered near the N-terminus of the

XX mature pVIII. The phages are antigenically-homogeneous such that they

XX are likely to induce a restricted antibody response. The phages are

XX easily produced and purified. They are also easy to engineer by

CC recombinant methods and the large size of the phage may lend to its
CC adjuvant-like effects.
XX
SQ Sequence 24 AA;

Query Match 100.0%; Score 115; DB 22; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPVDPNANPNANPNANP 20
Db 1 NANPVDPNANPNANPNANP 20

RESULT 6
AAU93822
ID AAU93822 standard; Peptide: 24 AA.
XX
AC AAU93822;

02-JUL-2002 (first entry)

DE P. falciparum CS B cell epitope #2.
XX
KW Immunogenic; hepatitis nucleocapsid protein; hepatitis B core; HBC;
KM vaccine; B cell epitope; T cell epitope; immunostimulant.
XX
OS Plasmodium falciparum.
XX
PN WO200214478-A2.
XX
PD 21-FEB-2002.
PE 16-AUG-2001; 2001WO-US41759.
PF 16-AUG-2000; 2000US-225843P.
PR 22-AUG-2000; 2000US-226867P.
PR 15-AUG-2001; 2001US-0930915.
XX
PA (APOV-) APOVIA INC.
XX
PI Birkett AJ;
XX
DR WPI: 2002-257601/30.
XX
PT Novel recombinant hepatitis nucleocapsid protein, termed as chimeric
PT hepatitis B core protein, displays immunogenic epitopes at N-terminus,
PT HBC immunogenic loop with linker for conjugated epitope and C-terminus

PS Example 4; Page 121; 289pp; English.
XX
XX The invention relates to a recombinant hepatitis nucleocapsid protein,
CC i.e. a chimeric hepatitis B core (HBC) protein (I), displaying one or
CC more immunogenic epitopes at the N-terminus, HBC immunogenic loop (L) or
CC C-terminus, or having a heterologous linker for a conjugated epitope in
CC (L), and containing a Cys residue at, or near, the C-terminus that
CC confers enhanced stability to the particles. A vaccine comprising (I) is
CC useful for inducing an immune response in an inoculated host animal, by
CC inoculating a host animal with the vaccine, and maintaining that
CC inoculated animal for a time period sufficient for that animal to
CC develop an immune response. The immunogenic particles formed using (I)
CC are substantially free of binding to nucleic acids, and are most stable
CC than the particle formed from otherwise identical HBC chimera that lacks
CC the C-terminal residue or in which a C-terminal Cys is replaced by
CC another residue. The chimera particles are most stable on storage in
CC aqueous compositions that are particles of similar sequence that lack any
CC C-terminal Cys residues. The chimera molecule exhibits the self-assembly
CC not exhibiting the nucleic acid binding of those native particles, and
CC excellent B cell and T cell immunogenicities. The chimera particles are
CC typically prepared in higher yield than similar particles that are free
CC of a C-terminal Cys. The particles are often far more immunogenic than
CC the similar conjugates that lack a C-terminal Cys. Immunogenicities of

CC particles assembled from the chimera molecules are enhanced as compared to
CC similar particles assembled from chimera molecules lacking at least one
CC C-terminal Cys. AAU93802-AAU93997 represent immunogenic HBC particles
CC amino acid sequences and related sequences of the invention.
XX
SQ Sequence 24 AA;

Query Match 100.0%; Score 115; DB 23; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPVDPNANPNANPNANP 20
Db 1 NANPVDPNANPNANPNANP 20

RESULT 7
AAU87696
ID AAU87696 standard; Peptide: 24 AA.
XX
AC AAU87696;
XX
DT 21-MAY-2002 (first entry)
XX
DE Malaria B cell epitope #2.
XX
KW Hepatitis B virus; nucleocapsid protein; HBC; hepatitis B virus core;
KM B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoa;
KM circumsporozoite; human immunodeficiency virus type I; human; squirrel;
KM woodchuck.
XX
OS Plasmodium falciparum.
XX
PN WO200213765-A2.
XX
PD 21-FEB-2002.
PE 16-AUG-2001; 2001WO-US25625.
PF 16-AUG-2000; 2000US-225813P.
PR 15-AUG-2001; 2001US-0931325.
XX
PA (APOV-) APOVIA INC.
XX
PI Birkett AJ;
XX
DR WPI: 2002-241832/29.
XX
PT Recombinant hepatitis B virus core protein chimera molecule, useful to
PT induce antibodies to malarial parasites, comprises malaria-specific
PT T-cell epitope and is engineered for enhanced stability -

PS Claim 49; Page 97; 197pp; English.
XX
XX The invention relates to a recombinant hepatitis B virus core (HBC)
CC protein chimera molecule that contains 4 peptide-linked amino acid residue
CC sequence domains. The molecule of the invention contains a region
CC constituting a B cell epitope of the circumsporozoite protein of a
CC species of the parasite, Plasmodium. The chimera sequence is useful as an
CC immunogen for inducing antibodies to the malaria-causing parasite,
CC Plasmodium, particularly P. falciparum and P. vivax. Sequences
CC AAU87695-AAU87804 represent peptide epitopes of the invention.
XX
SQ Sequence 24 AA;

Query Match 100.0%; Score 115; DB 23; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPVDPNANPNANPNANP 20
Db 1 NANPVDPNANPNANPNANP 20


```

RESULT 8
AAU93825
ID AAU93825 standard; Peptide; 28 AA.
XX
AC AAU93825;
XX
DT 02-JUL-2002 (first entry)
XX
DE P. falciparum CS B cell epitope #5.
XX
KW Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; HBC;
KM vaccine; B cell epitope; T cell epitope; immunostimulant.
XX
OS Plasmodium falciparum.
XX
PN WO200214478-A2.
XX
PD 21-FEB-2002.
XX
PR 16-AUG-2001; 2001WO-US41759.
XX
PR 16-AUG-2000; 2000US-225843P.
XX
PR 22-AUG-2000; 2000US-226867P.
XX
PR 15-AUG-2001; 2001US-0930915.
XX
PA (APOV-) APOVIA INC.
XX
PI Birkett AJ;
XX
DR WPI; 2002-257601/30.
XX
PT Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric
PT hepatitis B core protein, displays immunogenic epitopes at N-terminus,
PT HBC immunogenic loop with linker for conjugated epitope and C-terminus
PT
XX
PS Example 4; Page 121; 289pp; English.
XX
CC The invention relates to a recombinant hepadnavirus nucleocapsid protein,
CC i.e. a chimeric hepatitis B core (HBC) protein (I), displaying one or
CC more immunogenic epitopes at the N-terminus, HBC immunogenic loop (L) or
CC C-terminus, or having a heterologous linker for a conjugated epitope in
CC (L), and containing a Cys residue at, or near, the C-terminus that
CC confers enhanced stability to the particles. A vaccine comprising (I) is
CC useful for inducing an immune response in an inoculated host animal, by
CC inoculating a host animal with the vaccine, and maintaining that
CC inoculated animal for a time period sufficient for that animal to
CC develop an immune response. The immunogenic particles formed using (I)
CC are substantially free of binding to nucleic acids, and are most stable
CC than the particle formed from otherwise identical HBC chimera that lacks
CC the C-terminal residue or in which a C-terminal Cys is replaced by
CC another residue. The chimera particles are most stable on storage in
CC aqueous compositions that are particles of similar sequence that lack any
CC C-terminal Cys residues. The chimera molecule exhibits the self-assembly
CC not exhibiting the nucleic acid binding of those native particles, and
CC excellent B cell and T cell immunogenicities. The chimera particles are
CC typically prepared in higher yield than similar particles that are free
CC of a C-terminal Cys. The particles are often far more immunogenic than
CC the similar conjugates that lack a C-terminal Cys. Immunogenicities of
CC particles assembled from the chimera molecules are enhanced as compared to
CC similar particles assembled from chimera molecules lacking at least one
CC C-terminal Cys. AAU93802-AAU93997 represent immunogenic HBC particles
CC amino acid sequences and related sequences of the invention.
XX
SQ Sequence 28 AA;

```

```

Db 1 NANPVDPNANPNANPNANP 20
RESULT 9
AAU87699
ID AAU87699 standard; Peptide; 28 AA.
XX
AC AAU87699;
XX
DT 21-MAY-2002 (first entry)
XX
DE Malarial B cell epitope #5.
XX
KW Hepatitis B virus; nucleocapsid protein; HBC; hepatitis B virus core;
KM B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoacide;
KM circumsporozoite; human immunodeficiency virus type I; human; squirrel;
KM woodchuck.
XX
OS Plasmodium falciparum.
XX
PN WO200213765-A2.
XX
PD 21-FEB-2002.
XX
PR 16-AUG-2001; 2001WO-US25625.
XX
PR 16-AUG-2000; 2000US-225813P.
XX
PR 15-AUG-2001; 2001US-0931325.
XX
PA (APOV-) APOVIA INC.
XX
PI Birkett AJ;
XX
DR WPI; 2002-241832/29.
XX
PT Recombinant hepatitis B virus core protein chimera molecule, useful to
PT induce antibodies to malarial parasites, comprises malaria-specific
PT T-cell epitope and is engineered for enhanced stability -
XX
XX
PS Claim 49; Page 31; 197pp; English.
XX
CC The invention relates to a recombinant hepatitis B virus core (HBC)
CC protein chimera molecule that contains 4 peptide-linked amino acid residue
CC sequence domains. The molecule of the invention contains a region
CC constituting a B cell epitope of the circumsporozoite protein of a
CC species of the parasite, Plasmodium. The chimera sequence is useful as an
CC immunogen for inducing antibodies to the malaria-causing parasite,
CC Plasmodium, particularly P. falciparum and P. vivax. Sequences
CC AAU87695-AAU87804 represent peptide epitopes of the invention.
XX
SQ Sequence 28 AA;
Query Match 100.0%; Score 115; DB 23; Length 28;
Best Local Similarity 100.0%; Pred. No. 4.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NANPVDPNANPNANPNANP 20
Db 1 NANPVDPNANPNANPNANP 20
RESULT 10
AAU93867
ID AAU93867 standard; Peptide; 31 AA.
XX
AC AAU93867;
XX
DT 02-JUL-2002 (first entry)
XX
DE P. falciparum CS-repeat B cell epitope PF3.
XX
KW Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; HBC;
KM vaccine; B cell epitope; T cell epitope; immunostimulant.

```

XX Plasmodium falciparum.
 OS
 XX
 PN WO200214478-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 16-AUG-2001; 2001WO-USA1759.
 XX
 PR 16-AUG-2000; 2000US-225843P.
 PR 22-AUG-2000; 2000US-226867P.
 PR 15-AUG-2001; 2001US-0930915.
 XX
 PA (APOV-) APOVIA INC.
 XX
 PI Birkett AJ;
 XX
 DR WPI; 2002-257601/30.

Novel recombinant hepatitis B core protein, displays immunogenic epitopes at N-terminus, Hbc immunogenic loop with linker for conjugated epitope and C-terminus

PT
 XX
 PS Example 1; Page 106; 289pp; English.

CC The invention relates to a recombinant hepatitis virus nucleocapsid protein, i.e. a chimeric hepatitis B core (Hbc) protein (I), displaying one or more immunogenic epitopes at the N-terminus, Hbc immunogenic loop (L) or C-terminus, or having a heterologous linker for a conjugated epitope in (L) and containing a Cys residue at, or near, the C-terminus that confers enhanced stability to the particles. A vaccine comprising (I) is useful for inducing an immune response in an inoculated host animal, by inoculating a host animal with the vaccine, and maintaining that inoculated animal for a time period sufficient for that animal to develop an immune response. The immunogenic particles formed using (I) are substantially free of binding to nucleic acids, and are most stable than the particle formed from otherwise identical Hbc chimera that lacks the C-terminal residue or in which a C-terminal Cys is replaced by another residue. The chimera particles are most stable on storage in aqueous compositions that are particles of similar sequence that lack any C-terminal Cys residues. The chimera molecule exhibits the self-assembly not exhibiting the nucleic acid binding of those native particles, and excellent B cell and T cell immunogenicities. The chimera particles are typically prepared in higher yield than similar particles that are free of a C-terminal Cys. The particles are often far more immunogenic than the similar conjugates that lack a C-terminal Cys. Immunogenicities of particles assembled from the chimera molecules are enhanced as compared to similar particles assembled from chimera molecules lacking at least one C-terminal Cys. AAU93802-AAU93997 represent immunogenic Hbc particles amino acid sequences and related sequences of the invention.

XX
 SQ Sequence 31 AA;

Query Match 100.0%; Score 115; DB 23; Length 31;

Best Local Similarity 100.0%; Pred. No. 5e-08; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPVPDPNANPNANPNANP 20
 ||||||||||||||||||
 Db 2 NANPVPDPNANPNANPNANP 21

RESULT 11

AAU93870
 ID AAU93870 standard; Peptide; 31 AA.

XX
 AC AAU93870;

DT 02-JUL-2002 (first entry)

XX P. falciparum CS-repeat B cell epitope Pf3.3.
 DE
 XX

KW Immunogenic; hepatitis virus nucleocapsid protein; hepatitis B core; Hbc; vaccine; B cell epitope; T cell epitope; immunostimulant.

OS Plasmodium falciparum.

XX WO200214478-A2.

XX PD 21-FEB-2002.

XX PF 16-AUG-2001; 2001WO-USA1759.

XX PR 16-AUG-2000; 2000US-225843P.

XX PR 22-AUG-2000; 2000US-226867P.

XX PR 15-AUG-2001; 2001US-0930915.

XX PA (APOV-) APOVIA INC.

XX PI Birkett AJ;

XX DR WPI; 2002-257601/30.

Novel recombinant hepatitis virus nucleocapsid protein, termed as chimeric hepatitis B core protein, displays immunogenic epitopes at N-terminus, Hbc immunogenic loop with linker for conjugated epitope and C-terminus

PT
 XX
 PS Example 1; Page 107; 289pp; English.

CC The invention relates to a recombinant hepatitis virus nucleocapsid protein, i.e. a chimeric hepatitis B core (Hbc) protein (I), displaying one or more immunogenic epitopes at the N-terminus, Hbc immunogenic loop (L) or C-terminus, or having a heterologous linker for a conjugated epitope in (L), and containing a Cys residue at, or near, the C-terminus that confers enhanced stability to the particles. A vaccine comprising (I) is useful for inducing an immune response in an inoculated host animal, by inoculating a host animal with the vaccine, and maintaining that inoculated animal for a time period sufficient for that animal to develop an immune response. The immunogenic particles formed using (I) are substantially free of binding to nucleic acids, and are most stable than the particle formed from otherwise identical Hbc chimera that lacks the C-terminal residue or in which a C-terminal Cys is replaced by another residue. The chimera particles are most stable on storage in aqueous compositions that are particles of similar sequence that lack any C-terminal Cys residues. The chimera molecule exhibits the self-assembly not exhibiting the nucleic acid binding of those native particles, and excellent B cell and T cell immunogenicities. The chimera particles are typically prepared in higher yield than similar particles that are free of a C-terminal Cys. The particles are often far more immunogenic than the similar conjugates that lack a C-terminal Cys. Immunogenicities of particles assembled from the chimera molecules are enhanced as compared to similar particles assembled from chimera molecules lacking at least one C-terminal Cys. AAU93802-AAU93997 represent immunogenic Hbc particles amino acid sequences and related sequences of the invention.

XX
 SQ Sequence 31 AA;

Query Match 100.0%; Score 115; DB 23; Length 31;

Best Local Similarity 100.0%; Pred. No. 5e-08; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPVPDPNANPNANPNANP 20
 ||||||||||||||||||
 Db 2 NANPVPDPNANPNANPNANP 21

RESULT 12

AAU87722
 ID AAU87722 standard; Peptide; 31 AA.

XX
 AC AAU87722;

DT 21-MAY-2002 (first entry)
 XX

DE Malarial B cell epitope #25.
 XX
 KW Hepatitis B virus; nucleocapsid protein; HBC; hepatitis B virus core;
 KW B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoa; acid;
 KW circumsporozoite; human immunodeficiency virus type I; human; squirrel;
 KW woodchuck.
 XX
 OS Plasmodium falciparum.
 XX
 PN W0200213765-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 16-AUG-2001; 2001WO-US25625.
 XX
 PR 16-AUG-2000; 2000US-225813P.
 XX
 PA 15-AUG-2001; 2001US-0931325.
 XX
 PA (APOV-) APOVIA INC.
 XX
 Birkett AJ;
 XX
 DR WPI, 2002-241832/29.
 DR N-PSDB; ABR44222.
 XX
 XX
 PT Recombinant hepatitis B virus core protein chimera molecule, useful to
 PT induce antibodies to malarial parasites, comprises malaria-specific
 PT T-cell epitope and is engineered for enhanced stability -
 XX
 PS Example 1; Page 84; 197pp; English.
 XX
 XX The invention relates to a recombinant hepatitis B virus core (HBC)
 CC protein chimera molecule that contains a peptide-linked amino acid residue
 CC sequence domains. The molecule of the invention contains a region
 CC constituting a B cell epitope of the circumsporozoite protein of a
 CC species of the parasite, Plasmodium. The chimera sequence is useful as an
 CC immunogen for inducing antibodies to the malaria-causing parasite,
 CC Plasmodium, particularly P. falciparum and P. vivax. Sequences
 CC AAU87695-AAU87804 represent peptide epitopes of the invention.
 CC
 SO Sequence 31 AA;
 XX
 Query Match 100.0%; Score 115; DB 23; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NANPNVDPNANPNANPNANP 20
 ||||||||||||||||||
 2 NANPNVDPNANPNANPNANP 21
 ||||||||||||||||||
 RESULT 13
 AAU87725
 ID AAU87725 standard; Peptide; 31 AA.
 XX
 AC AAU87725;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Malarial B cell epitope #28.
 XX
 KW Hepatitis B virus; nucleocapsid protein; HBC; hepatitis B virus core;
 KW B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoa; acid;
 KW circumsporozoite; human immunodeficiency virus type I; human; squirrel;
 KW woodchuck.
 XX
 OS Plasmodium falciparum.
 XX
 PN W0200213765-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 16-AUG-2001; 2001WO-US25625.

XX
 PR 16-AUG-2000; 2000US-225813P.
 PR 15-AUG-2001; 2001US-0931325.
 XX
 PA (APOV-) APOVIA INC.
 XX
 PI Birkett AJ;
 XX
 DR WPI, 2002-241832/29.
 DR N-PSDB; ABR44228.
 XX
 XX
 PT Recombinant hepatitis B virus core protein chimera molecule, useful to
 PT induce antibodies to malarial parasites, comprises malaria-specific
 PT T-cell epitope and is engineered for enhanced stability -
 XX
 PS Example 1; Page 85; 197pp; English.
 XX
 XX The invention relates to a recombinant hepatitis B virus core (HBC)
 CC protein chimera molecule that contains a peptide-linked amino acid residue
 CC sequence domains. The molecule of the invention contains a region
 CC constituting a B cell epitope of the circumsporozoite protein of a
 CC species of the parasite, Plasmodium. The chimera sequence is useful as an
 CC immunogen for inducing antibodies to the malaria-causing parasite,
 CC Plasmodium, particularly P. falciparum and P. vivax. Sequences
 CC AAU87695-AAU87804 represent peptide epitopes of the invention.
 CC
 SO Sequence 31 AA;
 XX
 Query Match 100.0%; Score 115; DB 23; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NANPNVDPNANPNANPNANP 20
 ||||||||||||||||||
 Db 2 NANPNVDPNANPNANPNANP 21
 ||||||||||||||||||
 RESULT 14
 AAP82566
 ID AAP82566 standard; protein; 38 AA.
 XX
 AC AAP82566;
 XX
 DT 19-NOV-1990 (first entry)
 XX
 DE Anti-malarial peptide.
 XX
 KW Malaria; antisporezoite antibodies.
 XX
 PN DE3741183-A.
 XX
 PD 09-JUN-1988.
 XX
 PF 04-DEC-1987; 87DE-3741183.
 XX
 PR 04-DEC-1986; 86IT-0022560.
 XX
 PA (ENIE) ENIRICERHE SPA.
 XX
 PI Bernardi A, Bonelli F, Pessi A, Verdini AS;
 XX
 DR WPI, 1988-162768/24.
 XX
 PT Immunologically active polypeptide - for prep. of anti-malaria
 PT vaccines and diagnostic appts. for determination of antisporezoite
 PT antibodies.
 XX
 PS Claim 1; Page 2; 7pp; German.
 XX
 CC This polypeptide can be used in the prep. of antimalarial vaccines
 CC and for detecting antisporezoite antibodies in clinical samples
 CC of malaria sufferers.
 XX

SQ Sequence 38 AA:

Query Match 100.0%; Score 115; DB 9; Length 38;

Best Local Similarity 100.0%; Pred. No. 6.1e-08; Mismatches 20; Conservative 0; Indels 0; Gaps 0;

QY 1 NANPVVDNANPNANPNANP 20

Db 13 NANPVVDNANPNANPNANP 32

RESULT 15

ID AAM59273 standard; peptide: 126 AA.

AC AAM59273;

DT 27-AUG-1998 (first entry)

P. falciparum circumsporozoite region II peptide fragment CS27IWC-His6.

Malaria; gene therapy; hepatocyte; liver; circumsporozoite; ligand; targeted delivery; therapy; disease; cancer; hepatitis; cystic fibrosis; hypercholesterolaemia; phenylketonuria; haemophilia.

Plasmodium falciparum.

PN US5766899-A.

PD 16-JUN-1998.

PF 27-FEB-1995; 95US-0395602.

PR 27-FEB-1995; 95US-0395602.

PA (TEXA) UNIV TEXAS SYSTEM.

PI Ding Z, Kuo MT;

WP; 1998-361692/31.

Complexes for targeted delivery of nucleic acids to hepatocytes -
containing Plasmodium circumsporozoite polypeptide as targeting
ligand

Claim 2; Column 31-32; 34pp; English.

AAM59270-W59274 are fragments of malarial circumsporozoite (CS) region
II isolated from Plasmodium falciparum. These fragments can be used as
ligands in a method for the targeted delivery of nucleic acid to cells
in culture or cells in vivo, especially where the cells are hepatocytes.
Therapy of diseases such as cancer, malaria, hepatitis, cystic fibrosis,
hypercholesterolaemia, phenylketonuria and haemophilia is mentioned.
CS polypeptides are liver cell specific with rapid hepatic invasion.
They are more efficient than the prior art asialoosomucoid (ASOR)
ligands, of which there may be an accumulation in certain diseases due to
receptor downregulation.

SQ Sequence 126 AA:

Query Match 100.0%; Score 115; DB 19; Length 126;

Best Local Similarity 100.0%; Pred. No. 2e-07; Mismatches 20; Conservative 0; Indels 0; Gaps 0;

QY 1 NANPVVDNANPNANPNANP 20

Db 17 NANPVVDNANPNANPNANP 36

Search completed: December 6, 2002, 12:43:35
Job time : 29.8525 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: December 6, 2002, 12:42:06 : Search time 11.4754 Seconds
(Without alignments)
51.280 Million cell updates/sec

Title: US-09-931-325C-3

Sequence: 1 NANPNVDPNANPNANPNANP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

tal number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/6CTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	126	1 US-08-395-602A-4	Sequence 4, Appl1
2	115	100.0	126	2 US-08-021-625D-4	Sequence 4, Appl1
3	115	100.0	163	5 PCT-US93-08435-2	Sequence 2, Appl1
4	115	100.0	412	1 US-08-313-288B-18	Sequence 18, Appl1
5	106	92.2	24	6 5219987-4	Patent No. 5219987
6	106	92.2	46	1 US-08-395-602A-5	Sequence 5, Appl1
7	106	92.2	46	2 US-08-021-625D-5	Sequence 5, Appl1
8	106	92.2	66	2 US-08-455-625-35	Sequence 35, Appl1
9	106	92.2	66	4 US-08-455-625-35	Sequence 35, Appl1
10	106	92.2	66	4 US-08-060-988A-35	Sequence 35, Appl1
11	106	92.2	66	5 PCT-US94-05142-35	Sequence 35, Appl1
12	106	92.2	68	1 US-08-143-365A-13	Sequence 13, Appl1
13	106	92.2	126	1 US-08-268-348A-12	Sequence 12, Appl1
14	106	92.2	133	1 US-08-268-348A-8	Sequence 8, Appl1
15	106	92.2	133	1 US-08-268-348A-10	Sequence 10, Appl1
16	106	92.2	160	6 PCT-US93-08435-1	Patent No. 5219987
17	106	92.2	164	5 PCT-US93-08435-1	Sequence 1, Appl1
18	106	92.2	423	2 US-08-760-797A-1	Sequence 1, Appl1
19	106	92.2	424	4 US-08-760-797A-3	Sequence 3, Appl1
20	106	92.2	424	4 US-08-932-929B-1	Sequence 1, Appl1
21	106	92.2	424	4 US-08-932-929B-3	Sequence 3, Appl1
22	83	72.2	17	1 US-07-987-286-8	Sequence 8, Appl1
23	83	72.2	17	2 US-08-614-626-8	Sequence 8, Appl1
24	70	60.9	28	3 US-08-822-324-23	Sequence 23, Appl1
25	70	60.9	1687	2 US-08-570-311-29	Sequence 29, Appl1
26	69	60.0	12	4 US-09-248-588-105	Sequence 105, Appl1
27	69	60.0	12	4 US-08-909-551-4	Sequence 4, Appl1

28	69	60.0	12	4 US-09-314-135-1	Sequence 1, Appl1
29	69	60.0	13	4 US-08-909-551-3	Sequence 3, Appl1
30	64	55.7	12	2 US-08-747-137-41	Sequence 41, Appl1
31	58	50.4	229	4 US-09-069-023-20	Sequence 20, Appl1
32	58	50.4	1704	3 US-08-336-308A-10	Sequence 10, Appl1
33	58	50.4	1704	3 US-08-822-324-6	Sequence 6, Appl1
34	58	50.4	1704	4 US-09-490-931-10	Sequence 10, Appl1
35	58	50.4	1732	2 US-08-570-311-10	Sequence 10, Appl1
36	58	50.4	1732	2 US-08-353-485-10	Sequence 10, Appl1
37	57	49.6	51	1 US-08-361-920-12	Sequence 12, Appl1
38	57	49.6	51	1 US-08-479-939-12	Sequence 12, Appl1
39	57	49.6	51	1 US-08-483-432-12	Sequence 12, Appl1
40	56	48.7	1087	2 US-08-570-311-8	Sequence 8, Appl1
41	56	48.7	1087	2 US-08-353-485-8	Sequence 8, Appl1
42	56	48.7	1358	2 US-08-570-311-27	Sequence 27, Appl1
43	56	48.7	2628	2 US-08-570-311-14	Sequence 14, Appl1
44	54	47.0	430	2 US-08-712-709-9	Sequence 9, Appl1
45	54	47.0	430	3 US-09-111-444-9	Sequence 9, Appl1

ALIGNMENTS

RESULT 1
US-08-395-602A-4

; Sequence 4, Application US/08395602A

; Patent No. 5766899

; GENERAL INFORMATION:

; APPLICANT: Kuo, M. Tien

; TITLE OF INVENTION: Targeted Nucleic Acid Delivery into

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESS: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: United States of America

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/395,602A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Highlander, Steven L.

; REGISTRATION NUMBER: 37,642

; REFERENCE/DOCKET NUMBER: UTSC:410/HYL

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000

; TELEFAX: (713) 789-2679

; TELEX: 79-0924

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 126 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; US-08-395-602A-4

Query Match 100.0%; Score 115; DB 1; Length 126;

Best Local Similarity 100.0%; Pred. No. 9.5e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNVDPNANPNANPNANP 20

Db 17 NANPNVDPNANPNANPNANP 36

RESULT 2
US-08-021-625D-4
Sequence 4, Application US/08021625D
Patent No. 5976851
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
APPLICANT: Ding, Zhi-Ming
TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into
TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021.625D
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-021-625D-4

Query Match 100.0%; Score 115; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 9.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPVDPNANPNANPNANP 20
|||||
17 NANPVDPNANPNANPNANP 36

RESULT 3
PCT-US93-08435-2
Sequence 2, Application PC/TUS9308435
GENERAL INFORMATION:
APPLICANT: SmithKline Beecham, Corporation
APPLICANT: U. S. Government, Secretary of
APPLICANT: the Navy
APPLICANT: U. S. Government, Secretary of
TITLE OF INVENTION: The Army
TITLE OF INVENTION: Novel Antibodies for Confering Passive
TITLE OF INVENTION: Immunity Against Infection by a Pathogen in Man
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Box 457, 321 Norristown Road
CITY: Spring House
STATE: PA
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08435
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/941,654
FILING DATE: 09-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Maury E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: SBC P50107
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 540-9200
TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 163 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US93-08435-2

Query Match 100.0%; Score 115; DB 5; Length 163;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPVDPNANPNANPNANP 20
Db 60 NANPVDPNANPNANPNANP 79

RESULT 4
US-08-313-288B-18
Sequence 18, Application US/08313288B
Patent No. 5750502
GENERAL INFORMATION:
APPLICANT: Jersell, Thomas M. and Avihu Klar
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313.288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-313-288B-18

Query Match 100.0%; Score 115; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPVDNPNANPNANPNANP 20
Db 204 NANPVDNPNANPNANPNANP 223

RESULT 5

5219987-4
Patent No. 5219987
APPLICANT: VERDINI, ANTONIO S.; PESSI, ANTONELLO; BONELLI,
FABIO
TITLE OF INVENTION: SEQUENTIAL POLYPEPTIDES ENDOWED WITH
IMMUNOLOGICAL ACTIVITY
NUMBER OF SEQUENCES: 7
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/227, 364
FILING DATE: 02-AUG-1988
SEQ ID NO: 4
LENGTH: 24
5219987-4

Query Match 92.2%; Score 106; DB 6; Length 24;
Best Local Similarity 90.0%; Pred. No. 2.3e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NANPVDNPNANPNANPNANP 20
Db 3 NANPVDNPNANPNANPNANP 22

RESULT 6

US-08-395-602A-5
Sequence 5, Application US/08395602A
Patent No. 5766899
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,602A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-395-602A-5

Query Match 92.2%; Score 106; DB 1; Length 46;
Best Local Similarity 90.0%; Pred. No. 4.8e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NANPVDNPNANPNANPNANP 20
Db 1 NANPVDNPNANPNANPNANP 20

RESULT 7

US-08-021-625D-5
Sequence 5, Application US/08021625D
Patent No. 5976851
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,625D
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-021-625D-5

Query Match 92.2%; Score 106; DB 2; Length 46;
Best Local Similarity 90.0%; Pred. No. 4.8e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NANPVDNPNANPNANPNANP 20
Db 1 NANPVDNPNANPNANPNANP 20

RESULT 8

US-08-455-625-35
Sequence 35, Application US/08455625
Patent No. 5932218
GENERAL INFORMATION:
APPLICANT: Berzoiaky, Jay A.
APPLICANT: Ahlers, Jeffrey D.
APPLICANT: Pendleton, C. D.

```

: APPLICANT: Nara, Peter
: APPLICANT: Shirai, Mutsunori
: TITLE OF INVENTION: COMPOSITE SYNTHETIC PEPTIDE CONSTRUCT
: TITLE OF INVENTION: ELICITING NEUTRALIZING ANTIBODIES AND CYTOTOXIC T
: TITLE OF INVENTION: LYMPHOCYTES AGAINST HIV
: NUMBER OF SEQUENCES: 36
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolasch & Birch
: STREET: P.O. Box 747
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22040-0747
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/455,625
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/060,988
: FILING DATE: 14-MAY-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Svensson, Leonard R.
: REGISTRATION NUMBER: 30330
: REFERENCE/DOCKET NUMBER: 1173-434P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-205-8000
: TELEFAX: 703-205-8050
: INFORMATION FOR SEQ ID NO: 35:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 66 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FRAGMENT TYPE: internal
: FEATURE:
: NAME/KEY: Peptide
: LOCATION: 1..66
: OTHER INFORMATION: /label= peptide
: OTHER INFORMATION: /note= "peptide from P. falciparum CS antigen"
: US-08-455-625-35
:
: Query Match 92.2%; Score 106; DB 2; Length 66;
: Best Local Similarity 90.0%; Pred. No. 7.2e-08;
: Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
:
: QY 1 NANPVDPNANPNANPNANP 20
: 11111111111111111111
: DB 22 NANPNANPNANPNANPNANP 41
:
: RESULT 9
: US-08-455-685-35
: Sequence 35, Application US/08455685
: Patent No. 6214347
: GENERAL INFORMATION:
: APPLICANT: Berzofsky, Jay A.
: APPLICANT: Ahlers, Jeffrey D.
: APPLICANT: Pendleton, C. David
: APPLICANT: Nara, Peter
: APPLICANT: Shirai, Mutsunori
: TITLE OF INVENTION: MULTIDETERMINANT PEPTIDES THAT ELICIT
: TITLE OF INVENTION: HELPER T-LYMPHOCYTE, CYTOTOXIC T LYMPHOCYTE AND
: TITLE OF INVENTION: NEUTRALIZING ANTIBODY RESPONSES AGAINST HIV-1
: NUMBER OF SEQUENCES: 40
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston

```

```

: STATE: MA
: COUNTRY: US
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows95
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/455,685
: FILING DATE: 31-MAY-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/060,988
: FILING DATE: 14-MAY-1993
: APPLICATION NUMBER: 07/847,311
: FILING DATE: 06-MAR-1992
: APPLICATION NUMBER: 07/751,998
: FILING DATE: 29-AUG-1991
: APPLICATION NUMBER: 07/148,692
: FILING DATE: 26-JAN-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Beattie, Ingrid A.
: REGISTRATION NUMBER: P-42,306
: REFERENCE/DOCKET NUMBER: 08830/022003
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 35:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 66 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-455-685-35
:
: Query Match 92.2%; Score 106; DB 4; Length 66;
: Best Local Similarity 90.0%; Pred. No. 7.2e-08;
: Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
:
: QY 1 NANPVDPNANPNANPNANP 20
: 11111111111111111111
: DB 22 NANPNANPNANPNANPNANP 41
:
: RESULT 10
: US-08-060-988A-35
: Sequence 35, Application US/08060988A
: Patent No. 6294322
: GENERAL INFORMATION:
: APPLICANT: Berzofsky, Jay A.
: APPLICANT: Ahlers, Jeffrey D.
: APPLICANT: Pendleton, C. David
: APPLICANT: Nara, Peter
: APPLICANT: Shirai, Mutsunori
: TITLE OF INVENTION: MULTIDETERMINANT PEPTIDES
: TITLE OF INVENTION: THAT ELICIT
: TITLE OF INVENTION: HELPER T-LYMPHOCYTE, CYTOTOXIC T LYMPHOCYTE AND
: TITLE OF INVENTION: NEUTRALIZING ANTIBODY RESPONSES AGAINST HIV-1
: NUMBER OF SEQUENCES: 48
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: US
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows95
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:

```


APPLICATION NUMBER: US/08/060,988A
FILING DATE: 14-MAY-1993
PRIOR APPLICATION DATA: 07/847,311
APPLICATION NUMBER: 06-MAR-1992
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/751,998
FILING DATE: 29-AUG-1991
APPLICATION NUMBER: 07/148,692
FILING DATE: 26-JAN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Beattie, Ingrid A.
REGISTRATION NUMBER: P-42,306
REFERENCE/DOCKET NUMBER: 08830/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-060-988A-35

Query Match 92.2%; Score 106; DB 4; Length 66;
Best Local Similarity 90.0%; Pred. No. 7.2e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NANPNVDPNANPNANPNANP 20
Db 22 NANPNANPNANPNANPNANP 41

RESULT 11
PCT-US94-05142-35
Sequence 35, Application PC/TUS9405142
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: COMPOSITE SYNTHETIC PEPTIDE CONSTRUCT
TITLE OF INVENTION: ELICITING NEUTRALIZING ANTIBODIES AND CYTOTOXIC T
TITLE OF INVENTION: LYMPHOCYTES AGAINST HIV
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05142
FILING DATE: 13-MAY-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/060,988
FILING DATE: 14-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Svensson, Leonard R.
REGISTRATION/DOCKET NUMBER: 30330
REFERENCE/DOCKET NUMBER: 1173-434P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..66
OTHER INFORMATION: /label="peptide
PCT-US94-05142-35

Query Match 92.2%; Score 106; DB 5; Length 66;
Best Local Similarity 90.0%; Pred. No. 7.2e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NANPNVDPNANPNANPNANP 20
Db 22 NANPNANPNANPNANPNANP 41

RESULT 12
US-08-143-365A-13
Sequence 13, Application US/08143365A
Patent No. 5726292
GENERAL INFORMATION:
APPLICANT: Lowell, George H
TITLE OF INVENTION: IMMUNO-POTENTIATING SYSTEMS FOR
TITLE OF INVENTION: PREPARATION OF IMMUNOGENIC MATERIALS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hendricks and Assoc.
STREET: 9669 A Main Street
CITY: Fairfax
STATE: VA
COUNTRY: US
ZIP: 22031
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/143,365A
FILING DATE: 29-OCT-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Hendricks, Glenna M
REGISTRATION NUMBER: 32,535
REFERENCE/DOCKET NUMBER: Lowell-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 425-4250
TELEFAX: (703) 425-2767
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-143-365A-13

Query Match 92.2%; Score 106; DB 1; Length 68;
Best Local Similarity 90.0%; Pred. No. 7.4e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NANPNVDPNANPNANPNANP 20
Db 1 NANPNANPNANPNANPNANP 20

RESULT 13

US-08-268-348A-12
: Sequence 12, Application US/08268348A
: Patent No. 5750374
: GENERAL INFORMATION:
: APPLICANT: Dobelli, Heinz
: APPLICANT: Draeger, Nicholas
: APPLICANT: Trotman, Gerda H
: APPLICANT: Jakob, Peter
: APPLICANT: Stuber, Dietrich
: TITLE OF INVENTION: Process for Producing Hydrophobic
: TITLE OF INVENTION: Polypeptides and Proteins, and Fusion Proteins for Use in
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hoffmann-La Roche Inc.
: STREET: 340 Kingsland Street
: CITY: Nutley
: STATE: New Jersey
: COUNTRY: U.S.A.
: ZIP: 07110
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/268,348A
: FILING DATE: 29-JUN-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: EP 93110755.1
: FILING DATE: 06-JUL-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Parise, John P.
: REGISTRATION NUMBER: 34,403
: REFERENCE/DOCKET NUMBER: 4105/157
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (201) 235-6326
: TELEFAX: (201) 235-3500
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 126 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-268-348A-12

Query Match 92.2%; Score 106; DB 1; Length 126;
Best Local Similarity 90.0%; Pred. No. 1.5e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NANPNVDPNANPNANPNANP 20
|||||:|||||
Db 61 NANPNANPNANPNANPNANP 80

RESULT 14
US-08-268-348A-8
: Sequence 8, Application US/08268348A
: Patent No. 5750374
: GENERAL INFORMATION:
: APPLICANT: Dobelli, Heinz
: APPLICANT: Draeger, Nicholas
: APPLICANT: Trotman, Gerda H
: APPLICANT: Jakob, Peter
: APPLICANT: Stuber, Dietrich
: TITLE OF INVENTION: Process for Producing Hydrophobic
: TITLE OF INVENTION: Polypeptides and Proteins, and Fusion Proteins for Use in
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hoffmann-La Roche Inc.
: STREET: 340 Kingsland Street

CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/268,348A
FILING DATE: 29-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 93110755.1
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Parise, John P.
REGISTRATION NUMBER: 34,403
REFERENCE/DOCKET NUMBER: 4105/157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-6326
TELEFAX: (201) 235-3500
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 133 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-268-348A-8

Query Match 92.2%; Score 106; DB 1; Length 133;
Best Local Similarity 90.0%; Pred. No. 1.6e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NANPNVDPNANPNANPNANP 20
|||||:|||||
Db 65 NANPNANPNANPNANPNANP 84

RESULT 15
US-08-268-348A-10
: Sequence 10, Application US/08268348A
: Patent No. 5750374
: GENERAL INFORMATION:
: APPLICANT: Dobelli, Heinz
: APPLICANT: Draeger, Nicholas
: APPLICANT: Trotman, Gerda H
: APPLICANT: Jakob, Peter
: APPLICANT: Stuber, Dietrich
: TITLE OF INVENTION: Process for Producing Hydrophobic
: TITLE OF INVENTION: Polypeptides and Proteins, and Fusion Proteins for Use in
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hoffmann-La Roche Inc.
: STREET: 340 Kingsland Street
: CITY: Nutley
: STATE: New Jersey
: COUNTRY: U.S.A.
: ZIP: 07110
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/268,348A
: FILING DATE: 29-JUN-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: EP 93110755.1
: FILING DATE: 06-JUL-1993

ATTORNEY/AGENT INFORMATION:
 NAME: Parise, John P.
 REGISTRATION NUMBER: 34,403
 REFERENCE/DOCKET NUMBER: 4105/157
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (201) 235-6326
 TELEFAX: (201) 235-3500
 INFORMATION FOR SEO ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 133 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-268-348A-10

Query Match 92.28; Score 106; DB 1; Length 133;
 Best Local Similarity 90.08; Pred. No. 1.6e-07;
 Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 NANPNVDPNANPNANPNANP 20
 DB 65 NANPNANPNANPNANPNANP 84

Search completed: December 6, 2002, 12:46:43
 Job time: 12.4754 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:44:06 ; Search time 6.22951 Seconds
(without alignments)
52.146 Million cell updates/sec

Title: US-09-931-325C-3

Perfect score: 115

Sequence: 1 NANPNVDPNANPNANPNANP 20

BLOSUM62

Gapop 10.0 , Gapept 0.5

Scoring table: 103943 segs, 16242309 residues

tal number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCUS07_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	164	9	US-10-024-860-1
2	106	92.2	68	9	US-09-938-406-5
3	58	50.4	2150	9	US-10-135-322-17
4	57	49.6	367	10	US-09-971-118-2
5	56	48.7	50	10	US-09-864-761-47822
6	54	47.0	430	10	US-09-810-808-9
7	54	47.0	431	9	US-09-981-353-7
8	54	47.0	431	10	US-09-810-808-5
9	53	46.1	470	9	US-10-006-950-2
10	53	46.1	470	10	US-09-805-467A-2
11	51	44.3	532	10	US-09-828-302-12
12	49	42.6	659	12	US-10-090-602-12
13	49	42.6	802	10	US-09-287-849-10
14	49	42.6	2368	10	US-09-815-242-5635
15	49	42.6	2368	10	US-09-815-242-12389
16	49	42.6	2478	10	US-09-815-242-5816
17	49	42.6	2478	10	US-09-815-242-12967
18	47.5	41.3	792	10	US-09-995-587A-11
19	46	40.0	10	12	US-10-042-202-24

20	46	40.0	349	10	US-09-780-996-7	Sequence 7, Appli
21	44.5	38.7	877	10	US-09-881-752A-28	Sequence 28, Appl
22	44.5	38.7	2076	10	US-09-815-242-5815	Sequence 5815, Ap
23	44.5	38.7	2186	10	US-09-815-242-12913	Sequence 12913, A
24	44	38.3	79	10	US-09-764-846-193	Sequence 193, App
25	44	38.3	106	10	US-09-925-299-11384	Sequence 1384, Ap
26	44	38.3	1607	10	US-09-938-275-10	Sequence 10, Appl
27	43.5	37.8	479	8	US-08-910-386A-24	Sequence 24, Appl
28	43	37.4	63	10	US-09-864-761-43631	Sequence 43631, A
29	43	37.4	118	10	US-09-925-300-1823	Sequence 1823, Ap
30	43	37.4	174	10	US-09-818-066-54	Sequence 54, Appl
31	43	37.4	912	10	US-09-905-983-2	Sequence 2, Appli
32	43	37.4	912	10	US-09-746-491-49	Sequence 49, Appl
33	42.5	37.0	704	10	US-09-801-368-218	Sequence 218, App
34	42	36.5	174	10	US-09-818-066-52	Sequence 52, Appl
35	42	36.5	174	10	US-09-818-066-53	Sequence 53, Appl
36	42	36.5	174	10	US-09-818-066-56	Sequence 56, Appl
37	42	36.5	276	9	US-09-764-868-811	Sequence 811, App
38	42	36.5	496	10	US-09-801-368-418	Sequence 418, App
39	42	36.5	566	10	US-09-801-196-36	Sequence 36, Appl
40	42	36.5	707	10	US-09-842-256-3	Sequence 3, Appli
41	42	36.5	1036	10	US-09-845-583-10	Sequence 10, Appl
42	42	36.5	1587	10	US-10-001-535-146	Sequence 146, App
43	41.5	36.1	743	10	US-09-771-161A-164	Sequence 164, App
44	41.5	36.1	743	10	US-09-771-161A-254	Sequence 254, App
45	41.5	36.1	743	10	US-09-771-161A-254	Sequence 254, App

ALIGNMENTS

```
RESULT 1
US-10-024-860-1
; Sequence 1, Application US/10024860
; Patent No. US20020172692A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Vaccine Composition Against Malaria
; FILE REFERENCE: B45088C2
; CURRENT APPLICATION NUMBER: US/10/024,860
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 09/826,513
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/230,629
; PRIOR FILING DATE: 1999-01-26
; PRIOR APPLICATION NUMBER: GB 9616351.4
; PRIOR FILING DATE: 1996-08-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-024-860-1

Query Match      100.0%; Score 115; DB 9; Length 164;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NANPNVDPNANPNANPNANP 20
      |||
Db      81 NANPNVDPNANPNANPNANP 100

RESULT 2
US-09-938-406-5
; Sequence 5, Application US/09938406
; Patent No. US20020155120A1
; GENERAL INFORMATION:
; APPLICANT: Lowell, George
; APPLICANT: Vancott, Thomas
; APPLICANT: Birx, Deborah
; TITLE OF INVENTION: PROTEIN AND PEPTIDE VACCINES FOR
```

```
; TITLE OF INVENTION: INDUCING MUCOSAL IMMUNITY
; FILE REFERENCE: 40646-20002.10
; CURRENT APPLICATION NUMBER: US/09/938,406
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/214,701
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/US 97/12253
; PRIOR FILING DATE: 1997-07-10
; PRIOR APPLICATION NUMBER: US 60/021,687
; PRIOR FILING DATE: 1996-07-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 68
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-938-406-5
```

```
Query Match          92.2%; Score 106; DB 9; Length 68;
Best Local Similarity 90.0%; Pred. No. 3,9e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 NANPVDPNANPNANP 20
    |||||:|||||
DB 1 NANPNANPNANPNANP 20
```

```
RESULT 3
US-10-135-322-17
; Sequence 17, Application US/10135322
; Patent No. US20020173017A1
; GENERAL INFORMATION:
; APPLICANT: BENEFY, PN
; APPLICANT: HELARIUTTA, Y
; APPLICANT: MAHONEN, AP
; APPLICANT: BONKE, AMW
; APPLICANT: KAUPPINEN, L
; APPLICANT: RIIKKONEN, M
; TITLE OF INVENTION: WOODEN LEG GENE, PROMOTER AND USES THEREOF
; FILE REFERENCE: 5914-086-999
; CURRENT APPLICATION NUMBER: US/10/135,322
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 60/253,739
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 2150
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-135-322-17
```

```
Query Match          50.4%; Score 58; DB 9; Length 2150;
Best Local Similarity 52.6%; Pred. No. 3.4;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
OY 1 NANPVDPNANPNANP 19
    |||||:|||||
DB 416 NNNNNINNNNNNSN 434
```

```
RESULT 4
US-09-971-118-2
; Sequence 2, Application US/09971118
; Patent No. US20020123056A1
; GENERAL INFORMATION:
; APPLICANT: DELANEY, ALLEN
; APPLICANT: YOGANATHAN, THILAINATHAN
; TITLE OF INVENTION: SGR2 AND ITS USES
; FILE REFERENCE: KINE025C1P
; CURRENT APPLICATION NUMBER: US/09/971,118
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: PCT/US01/21479
```

```
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 60/237,419
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homin sapiens
US-09-971-118-2
```

```
Query Match          49.6%; Score 57; DB 10; Length 367;
Best Local Similarity 62.5%; Pred. No. 0.66;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 1 NANPVDPNANPNANP 16
    |||||:|||||
DB 17 NGNINLGSNPNAP 32
```

```
RESULT 5
US-09-864-761-47822
; Sequence 47822, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47822
```

```

; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC025129.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.47
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
US-09-864-761-47822

Query Match      48.7%; Score 56; DB 10; Length 50;
Best Local Similarity 61.1%; Pred. No. 0.1;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY      2 ANPNVDPNANPNANP 19
        ||| | | | | | | | | |
DB      13 ANANTDSNANSNTSNAN 30

RESULT 6
; US-09-810-808-9
; Sequence 9, Application US/09810808
; Patent No. US20020042114A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; Guegler, Karl J.
; Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/810,808
; FILING DATE: 15-Mar-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/541,228
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 294637
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-810-808-9

Query Match      47.0%; Score 54; DB 10; Length 430;
Best Local Similarity 42.3%; Pred. No. 2;
Matches 11; Conservative 4; Mismatches 5; Indels 6; Gaps 1;

OY      1 NANPNVDP-----NANPNANPNANP 20
        ||||| | | | | | | | | | |

```

```

DB      70 NANSPPPSPSQINLGPSSNPNAKP 95

RESULT 7
; US-09-981-353-7
; Sequence 7, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO: 7
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 3819039CD1
US-09-981-353-7

Query Match      47.0%; Score 54; DB 9; Length 431;
Best Local Similarity 42.3%; Pred. No. 2;
Matches 11; Conservative 4; Mismatches 5; Indels 6; Gaps 1;

OY      1 NANPNVDP-----NANPNANPNANP 20
        ||||| | | | | | | | | | |
DB      70 NANSPPPSPSQINLGPSSNPNAKP 95

RESULT 8
; US-09-810-808-5
; Sequence 5, Application US/09810808
; Patent No. US20020042114A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; Guegler, Karl J.
; Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/810,808
; FILING DATE: 15-Mar-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/541,228
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid

```

```
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: <unknown>
CLONE: Consensus
SEQUENCE DESCRIPTION: SEQ ID NO: 5
US-09-810-808-5
```

```
Query Match          47.0%; Score 54; DB 10; Length 431;
Best Local Similarity 42.3%; Pred. No. 2;
Matches 11; Conservative 4; Mismatches 5; Indels 6; Gaps 1;
```

```
Oy      1  NNPVNDP-----NANPNANPNANP 20
      |||||  |  |  |  |  |  |  |  |  |
Db      70  NNPSPSPSPSQINLGSSNPHAKP 95
```

RESULT 9

```
US-10-006-950-2
Sequence 2, Application US/10006950
Patent No. US20020161216A1
```

```
GENERAL INFORMATION:
APPLICANT: Borowsky, Beth E.
APPLICANT: Bonini, James A.
TITLE OF INVENTION: DNA ENCODING ORPHAN SNOREP4 RECEPTOR
FILE REFERENCE: 58799
CURRENT APPLICATION NUMBER: US/10/006,950
CURRENT FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: US/09/266,407
PRIOR FILING DATE: 1999-03-10
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0 - Beta
SEQ ID NO 2
LENGTH: 470
TYPE: PRT
ORGANISM: Homo sapiens
US-10-006-950-2
```

```
Query Match          46.1%; Score 53; DB 9; Length 470;
Best Local Similarity 44.4%; Pred. No. 2.9;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
```

```
Oy      3  NPNVDPNANPNANPNANP 20
      |||  |  |  |  |  |  |  |  |
Db      361 NPTLQPRSDPTAQPOLNP 378
```

RESULT 10

```
US-09-805-467A-2
Sequence 2, Application US/09805467A
Patent No. US20020058259A1
```

```
GENERAL INFORMATION:
APPLICANT: Ramakrishnan, Shyam
TITLE OF INVENTION: Regulation of Human Lipoxin A4
FILE REFERENCE: 4974,00453
CURRENT APPLICATION NUMBER: US/09/805,467A
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/189,037
PRIOR FILING DATE: 2000-03-14
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 470
TYPE: PRT
ORGANISM: Homo sapiens
US-09-805-467A-2
```

```
Query Match          46.1%; Score 53; DB 10; Length 470;
Best Local Similarity 44.4%; Pred. No. 2.9;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
```

```
Oy      3  NPNVDPNANPNANPNANP 20
      |||  |  |  |  |  |  |  |  |
Db      361 NPTLQPRSDPTAQPOLNP 378
```

RESULT 11

```
US-09-828-302-12
Sequence 12, Application US/09828302
Patent No. US20020152502A1
```

```
GENERAL INFORMATION:
APPLICANT: COSTA E SILVA, OSWALDO DA
APPLICANT: VAN THIELEN, NOCHA
APPLICANT: CHEN, ROUYING
APPLICANT: ISHITANI, MANABU
TITLE OF INVENTION: PHOSPHATASE STRESS-RELATED PROTEINS AND METHODS OF USE
FILE REFERENCE: 16313-0029
CURRENT APPLICATION NUMBER: US/09/828,302
CURRENT FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 60/196,001
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 532
TYPE: PRT
ORGANISM: Physcomitrella patens
US-09-828-302-12
```

```
Query Match          44.3%; Score 51; DB 10; Length 532;
Best Local Similarity 58.8%; Pred. No. 6.1;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
Oy      3  NPNVDPNANPNANPNANP 19
      |||||  |  |  |  |  |  |
Db      165 NLNVDPGARNGNPLSN 181
```

RESULT 12

```
US-10-090-624-12
Sequence 12, Application US/10090624
Patent No. US2002013235A1
```

```
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, MIO
APPLICANT: SHIMOTO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA=6
CURRENT APPLICATION NUMBER: US/10/090,624
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 659
TYPE: PRT
ORGANISM: Thermococcus celer
US-10-090-624-12
```

```
Query Match          42.6%; Score 49; DB 12; Length 659;
Best Local Similarity 80.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy      11  NPNANPNANP 20
      |||  |||  |||  |||
Db      545 NPNPNPNPNP 554
```



```
RESULT 13
US-09-287-849-10
; Sequence 10, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, David C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/918,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1996-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1996-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentl Ver. 2.1
; SEQ ID NO 10
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tetra-fusion
US-09-287-849-10
Query Match 42.6%; Score 49; DB 10; Length 802;
Best Local Similarity 61.9%; Pred. No. 18;
Matches 13; Conservative 1; Mismatches 5; Indels 2; Gaps 2;

QY 1 NANNVDN-ANPNANPNANP 20
Db 557 NAQPG-DPNAAPPADPNAPP 576

RESULT 14
US-09-815-242-5635
Sequence 5635, Application US/09815242
Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
```

```
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5635
; LENGTH: 2368
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5635

Query Match 42.6%; Score 49; DB 10; Length 2368;
Best Local Similarity 52.4%; Pred. No. 57;
Matches 11; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 2 ANPNVDNAN--PNANPNANP 20
Db 53 ADANTQPNANAGQANPTAQP 73

RESULT 15
US-09-815-242-12389
Sequence 12389, Application US/09815242
Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12389
; LENGTH: 2368
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12389

Query Match 42.6%; Score 49; DB 10; Length 2368;
Best Local Similarity 52.4%; Pred. No. 57;
Matches 11; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 2 ANPNVDNAN--PNANPNANP 20
Db 53 ADANTQPNANAGQANPTAQP 73
```

Search completed: December 6, 2002, 12:54:03
Job time : 7.22951 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:41:41 ; Search time 11.8033 Seconds
(without alignments)
162.894 Million cell updates/sec

Title: US-09-931-325c-3

Sequence: 1 NANPNVDPNANPNANPNANP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Optical number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database :
1: PIR:73:*
2: PIR:1:*
3: PIR:2:*
4: PIR:4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	405	2 S05428	circumsporozoite p
2	115	100.0	412	1 OZ20AF	circumsporozoite p
3	115	100.0	424	2 A54533	circumsporozoite p
4	115	100.0	442	2 A54529	circumsporozoite p
5	111	96.5	388	2 A39756	circumsporozoite p
6	70	60.9	169	2 T41898	hypothetical prote
7	70	60.9	1526	2 S49763	gingipain R (EC 3.
8	68	59.1	1528	2 D85912	hypothetical prote
9	68	59.1	1571	2 C91068	hypothetical prote
10	67	58.3	493	2 D90587	hypothetical prote
11	64	55.7	1171	2 T13065	PIP82 protein - fr
12	63	54.8	387	2 A86322	TRK14.9 protein -
13	62	53.9	356	2 A96826	AIDA-10 [imported
14	62	53.9	1327	2 B90674	probable beta-bar
15	62	53.9	1349	2 E85524	tspp protein, prob
16	61	53.0	501	2 B81048	tspp protein, prob
17	61	53.0	517	2 G81060	tspp protein, prob
18	61	53.0	548	2 A81070	tspp protein, prob
19	60.5	52.6	189	2 B29795	circumsporozoite p
20	59	51.3	252	2 T45737	hypothetical prote
21	58	50.4	1704	2 A55426	gingipain R (EC 3.
22	58	50.4	1732	2 T30836	lysine-specific cy
23	58	50.4	2150	2 S71629	sensory transduct
24	57.5	50.0	171	2 A29795	circumsporozoite p
25	57.5	50.0	415	2 F89994	hypothetical prote
26	57.5	50.0	589	2 AB1151	internalin protein
27	57	49.6	332	1 OZ20MB	circumsporozoite p
28	57	49.6	348	1 OZ20BK	circumsporozoite p
29	57	49.6	539	2 C81805	tspp protein NMA17

30	56	48.7	81	2 S51745	hypothetical prote
31	56	48.7	501	2 T48336	hypothetical prote
32	56	48.7	953	2 A86351	hypothetical prote
33	56	48.7	2628	2 T28651	hemagglutinin A -
34	55.5	48.3	790	2 T12203	transcription fact
35	54	47.0	431	2 A48094	serum and glucocor
36	54	47.0	482	2 AG1147	P60 extracellular
37	54	47.0	484	2 A41487	protein P60 precu
38	53	46.1	407	2 S53900	hypothetical prote
39	52.5	45.7	291	2 T31592	hypothetical prote
40	52	45.2	132	2 T49536	hypothetical prote
41	52	45.2	672	2 S46086	RNA-binding protei
42	52	45.2	967	2 H86334	720H2.10 protein -
43	51	44.3	194	2 J34589	immunoreactive pro
44	51	44.3	335	2 J35542	replication associ
45	51	44.3	346	2 T01123	hypothetical prote

ALIGNMENTS

RESULT 1
S05428
circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate NF54)
C:Species: Plasmodium falciparum
C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000
C:Accession: S05428; A45527; I60657
R:Campbell, J.R.
Nucleic Acids Res. 17, 5854, 1989
A:Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate
A:Reference number: S05428; MUID:89345189; PMID:266895
A:Accession: S05428
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-405 <CAM>
A:Cross-references: EMBL:X15363
R:Campbell, J.R.; Marsh, K.; Newbold, C.I.
Mol. Biochem. Parasitol. 37, 275-280, 1989
A:Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell
Mol. Biochem. Parasitol. 35, 185-190, 1989
A:Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate
A:Reference number: A45527; MUID:8936498; PMID:2671723
A:Accession: A45527
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <CAS>
A:Cross-references: GB:M22982; GB:J04650; NID:q160168; PIDN:AAA29527.1; PID:q160169
R:Lockyer, M.O.; Marsh, K.; Newbold, C.I.
Mol. Biochem. Parasitol. 37, 275-280, 1989
A:Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell
A:Reference number: A60657; MUID:90114334; PMID:2481827
A:Accession: I60657
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 319-336; 354-373 <LOC>
A:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:329-383/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 115; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 3; le-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NANPNVDPNANPNANPNANP 20
Db 201 NANPNVDPNANPNANPNANP 220

RESULT 2
OZ20AF
circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate INTW22)
C:Species: Plasmodium falciparum
C>Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jun-2000
C:Accession: A03388
R:Dame, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W
Science 225, 593-599, 1984

```

A>Title: Structure of the gene encoding the immunodominant surface antigen on the sporozoite
A:Reference number: A03386; MUID:84250215; PMID:6204383
A:Accession: A03386
A:Molecule type: DNA
A:Residues: 1-412 <DAN>
A:Cross-references: GB:K02194; NID:g160160; PIDN:AAA29524.1; PID:g160161
A:Experimental source: clone 7c8
C:Comment: Residues 1-16 are the probable signal sequence.
C:Comment: There are 41 copies of a 4-residue repeating unit in the middle domain of the
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
E:336-390/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match          100.0%; Score 115; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPVDPNANPNANPNANP 20
   |||||||||||||||||||
Db 204 NANPVDPNANPNANPNANP 223

SULF 3
A54533
Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain T4, Thailand)
C:Species: Plasmodium falciparum
C:Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text change 09-Jun-2000
C:Accession: A54533
R:del Portillo, H.A.; Nussenzweig, R.S.; Enea, V.
Mol. Biochem. Parasitol. 24, 289-294, 1987
A>Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.
A:Reference number: A54533; MUID:87315205; PMID:3306373
A:Accession: A54533
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <DEL>
A:Cross-references: GB:M19752; NID:g160216; PIDN:AAA29555.1; PID:g160217
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
E:3348-402/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match          100.0%; Score 115; DB 2; Length 424;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPVDPNANPNANPNANP 20
   |||||||||||||||||||
Db 140 NANPVDPNANPNANPNANP 159

SULF 4
529
Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain Wellcome)
C:Species: Plasmodium falciparum
C:Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text change 09-Jun-2000
C:Accession: A54529
R:Lockyer, M.J.; Schwartz, R.T.
Mol. Biochem. Parasitol. 22, 101-108, 1987
A>Title: Strain variation in the circumsporozoite protein gene of Plasmodium falciparum
A:Reference number: A54529; MUID:87115616; PMID:3543671
A:Accession: A54529
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-442 <LOC>
A:Cross-references: GB:M15505; NID:g160214; PIDN:AAA29554.1; PID:g160215
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: tandem repeat
E:366-420/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match          100.0%; Score 115; DB 2; Length 442;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPVDPNANPNANPNANP 20
   |||||||||||||||||||

```

```

DB      142  NANPVPDPNANPNNANPNANP 161

RESULT 5
A39756
Circumsporozoite protein - Plasmodium reichenowi
C:Species: Plasmodium reichenowi
C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999
C:Accession: A39756
R:Lat, A.A.: Goldman, I.F.
J. Biol. Chem. 266, 6686-6689, 1991
A:Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malar
A:Reference number: A39756; MUID:91201303; PMID:2016283
A:Accession: A39756
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <LAL>
C:Cross-references: GB:M60972; NID:g160228; PIDN:AAA29561.1; PID:g160229
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:312-366/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match          96.5%; Score 111; DB 2; Length 388;
Best Local Similarity 95.0%; Pred. No. 1e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1  NANPVPDPNANPNNANPNANP 20
        |||||
DB      148  NANPVPDPNANPNNANPNANP 167

RESULT 6
T41898
hypochemical protein H1 - human herpesvirus 7 (strain J1) (fragment)
C:Species: human herpesvirus 7
A:Variety: strain J1
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000
C:Accession: T41898; T41997
R:Nicholas, J.
submitted to the EMBL Data Library, December 1995
A:Description: Determination and analysis of the complete nucleotide sequence of huma
A:Reference number: Z22022
A:Accession: T41898
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-169 <NIC>
A:Cross-references: EMBL:U43400; PIDN:AAC54658.1
A:Genetics: GNI
A:Accession: T41997
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-169 <NI2>
A:Cross-references: EMBL:U43400; PIDN:AAC54757.1
A:Genetics: GN2
C:Genetics: <GN1>
A:Gene: H1
A:Map position: 33-542
C:Genetics: <GN2>
A:Gene: H1
A:Map position: 139080-139589

Query Match          60.9%; Score 70; DB 2; Length 169;
Best Local Similarity 55.0%; Pred. No. 0.013;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY      1  NANPVPDPNANPNNANPNANP 20
        |||||
DB      11  NNPNPNPSSKPNPSPNP 30

RESULT 7
S49763
gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis (fragment)
C:Species: Porphyromonas gingivalis

```

C>Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 31-Mar-1997
C/Accession: S49763
R/Author: Opoku, J.; Muir, J.; Stanley, J.M.; Rangarajan, M.; Curtis, M.A.
Submitted to the EMBL Data Library, November 1994
A/Description: Cloning, sequence analysis and expression in *Escherichia coli* of prp1 of
A/Reference number: S49763
A/Accession: S49763
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1526 <ADU>
A/Cross-references: EMBL:X82680
C/Genetics:
A/Genes: prp1
C/Keywords: cysteine proteinase; hydrolase

Query Match 60.9%; Score 70; DB 2; Length 1526;
Best Local Similarity 70.6%; Pred. No. 0.14;
Matches 12; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Y 4 PNVDPNANPNANPNANP 20
DB 946 PNGTPNPNPNPNPNP 962

RESULT 8

hypothetical protein ypiA [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93)
C/Species: *Escherichia coli*
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: D85912
R/Perla, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: D85912
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1528 <STO>
A/Cross-references: GB:AE005174; NID:g12517083; PIDN:AA657760.1; GSPDB:GN00145; UWGP:Z35
C/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Genes: ypiA

Query Match 59.1%; Score 68; DB 2; Length 1528;
Best Local Similarity 55.0%; Pred. No. 0.27;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Y 1 NANPNVDPNANPNANPNANP 20
DB 1181 DVKNPDPNPNPNPNP 1200

RESULT 9

hypothetical protein ECs3515 [imported] - *Escherichia coli* (strain O157:H7, substrain R1)
C/Species: *Escherichia coli*
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C/Accession: C91068
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
A/Reference number: A9629; MUID:21156231; PMID:11258796
A/Accession: C91068
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1571 <HAY>
A/Cross-references: GB:BA000007; PIDN:BA836938.1; PID:g13362986; GSPDB:GN00154
C/Experimental source: strain O157:H7, substrain R1MD 0509952
C/Genetics:
A/Genes: ECs3515

Query Match 59.1%; Score 68; DB 2; Length 1571;
Best Local Similarity 55.0%; Pred. No. 0.27;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Y 1 NANPNVDPNANPNANPNANP 20
DB 1224 DVKNPDPNPNPNPNP 1243

RESULT 10

lipoprotein [imported] - *Mycoplasma pulmonis* (strain UAB CTIP)
C/Species: *Mycoplasma pulmonis*
C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C/Accession: D90587
R/Chambaud, I.; Heilig, R.; Ferris, S.; Barde, V.; Samson, D.; Galisson, F.; Moszer,
Nucleic Acids Res. 29, 2145-2153, 2001
A/Title: The complete genome sequence of the murine respiratory pathogen *Mycoplasma p*
A/Reference number: A95512; MUID:21267165; PMID:11353084
A/Accession: D90587
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-493 <KUR>
A/Cross-references: GB:AL445566; PID:g14090019; PIDN:CAC13777.1; GSPDB:GN00153
A/Experimental source: strain UAB CTIP
C/Genetics:
A/Genes: MYPU_6040
A/Genetic code: GSC3

Query Match 58.3%; Score 67; DB 2; Length 493;
Best Local Similarity 63.2%; Pred. No. 0.1;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Y 1 NANPNVDPNANPNANPNANP 19
DB 101 NANPNQANPNQNTNPQN 119

RESULT 11

PIPB2 protein - fruit fly (*Drosophila melanogaster*)
C/Species: *Drosophila melanogaster*
C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C/Accession: T13065
R/Suri, V.; Qian, Z.; Hall, J.C.; Kosbash, M.
Neuron 21, 225-234, 1998
A/Title: Evidence that TIM light response is relevant to light-induced phase shifts 1
A/Reference number: Z17597; MUID:98361331; PMID:9697866
A/Accession: T13065
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-1171 <SUR>
A/Cross-references: EMBL:AF067153; NID:g3171244; PID:g3171245; PIDN:AAC18395.1
A/Experimental source: strain Canton-S, photoreceptor
C/Genetics:
A/Cross-references: FlyBase:Fbg0024943

Query Match 55.7%; Score 64; DB 2; Length 1171;
Best Local Similarity 55.0%; Pred. No. 0.68;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Y 1 NANPNVDPNANPNANPNANP 20
DB 1105 NLNANNSNPSTPNPNP 1124

RESULT 12

FA614.9 protein - *Arabidopsis thaliana*
C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C/Accession: A86322
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hutzar, E.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marzalli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, A.; Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86322
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-387 <STO>
A:Cross-references: GB:AE005172; NID:96730703; PIDN:AEF27098.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 54.8%; Score 63; DB 2; Length 387;
Best Local Similarity 61.1%; Pred. No. 0.27;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 NPNDPNANPNANPNANP 20
| : ||||| || : |
Db 78 NTHDHPNANPNPNPNP 95

RESULT 13
A96826
T8K14.10 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A96826
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hutzar, E.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marzalli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A96826
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-356 <STO>
A:Cross-references: GB:AE005173; NID:94835761; PIDN:ADJ0228.1; GSPDB:GN00141
C:Genetics:
A:Gene: T8K14.10
A:Map position: 1

Query Match 53.9%; Score 62; DB 2; Length 356;
Best Local Similarity 43.3%; Pred. No. 0.34;
Matches 13; Conservative 6; Mismatches 1; Indels 10; Gaps 2;

OY 1 NANPNVDPN-----ANPN-----ANPNANP 20
: ||| : || : ||| : ||
Db 93 SSNPNSNPPESSSNPNPDPSSSNPNNSNP 122

RESULT 14
B90674
Aida-I adhesin-like protein [imported] - *Escherichia coli* (strain 0157:H7, substrain RIM
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: B90674
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gasawara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* 0157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: B90674
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1327 <NAV>
A:Cross-references: PIDN:BA833785.1; PID:913359819; GSPDB:GN00154
A:Experimental source: strain 0157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs0362

Query Match 53.9%; Score 62; DB 2; Length 1327;
Best Local Similarity 58.8%; Pred. No. 1.4;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 4 PNVDPNANPNANPNANP 20
| : || || || || |
Db 974 PNPEPNPNPEPNPNPTP 990

RESULT 15
E85524
probable beta-barrel outer membrane protein Z0402 [imported] - *Escherichia coli* (stra
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: E85524
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* 0157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E85524
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1349 <STO>
A:Cross-references: GB:AE005174; NID:912513130; PIDN:AGS4657.1; GSPDB:GN00145; UWGP: C:Experimental source: strain 0157:H7, substrain EDL933
C:Genetics:
A:Gene: Z0402

Query Match 53.9%; Score 62; DB 2; Length 1349;
Best Local Similarity 58.8%; Pred. No. 1.5;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 4 PNVDPNANPNANPNANP 20
| : || || || || |
Db 996 PNPEPNPNPEPNPNPTP 1012

Search completed: December 6, 2002, 12:46:02
Job time : 12.8033 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:38:55 ; Search time 6.55738 Seconds

(without alignments)
126.503 Million cell updates/sec

Title: US-09-931-325C-3
Perfect score: 115
Sequence: 1 NANPNVDPNANPNANPNANP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	315	1	CSP_PLAFL
2	115	100.0	397	1	CSP_PLAFL
3	115	100.0	412	1	CSP_PLAFL
4	115	100.0	424	1	CSP_PLAFL
5	115	100.0	442	1	CSP_PLAFL
6	111	96.5	388	1	CSP_PLAFL
7	57	49.6	339	1	CSP_PLAFL
8	57	49.6	347	1	CSP_PLAFL
9	56	48.7	2628	1	HAGA_PORGI
10	54	47.0	430	1	SGK_RAT
11	54	47.0	431	1	SGK_HUMAN
12	54	47.0	431	1	SGK_MOUSE
13	54	47.0	431	1	SGK_RABIT
14	54	47.0	484	1	YNH8_YEAST
15	53	46.1	407	1	YNH8_YEAST
16	52	45.2	672	1	YNH8_YEAST
17	51	44.3	527	1	YNH8_YEAST
18	50	43.5	162	1	CRA_PLAFL
19	50	43.5	162	1	CRA_PLAFL
20	50	43.5	162	1	CRA_PLAFL
21	50	43.5	162	1	CRA_PLAFL
22	49	42.6	325	1	MODD_MYCBO
23	49	42.6	325	1	MODD_MYCBO
24	49	42.6	325	1	MODD_MYCBO
25	49	42.6	325	1	MODD_MYCBO
26	49	42.6	325	1	MODD_MYCBO
27	49	42.6	325	1	MODD_MYCBO
28	49	42.6	325	1	MODD_MYCBO
29	48	41.7	189	1	YHCH_BACSU
30	48	41.7	189	1	YHCH_BACSU
31	47.5	41.3	1314	1	SWIL_YEAST
32	47	40.9	198	1	SEGC_HELPJ
33	47	40.9	280	1	P32_MYCCE

34	47	40.9	370	1	CTPT_PLAFL
35	47	40.9	785	1	YE15_CAEEL
36	47	40.9	855	1	GAF1_SCHPO
37	46	40.0	122	1	YPS2_PLEPO
38	46	40.0	248	1	DAT1_YEAST
39	46	40.0	381	1	MODD_MYCAV
40	46	40.0	448	1	AAC2_DICDI
41	46	40.0	463	1	PLSB_CARTI
42	46	40.0	558	1	ROL_HUMAN
43	45.5	39.6	447	1	AC11_CAEEL
44	45	39.1	182	1	YN66_HALNI
45	45	39.1	199	1	SEGC_HELPJ

ALIGNMENTS

```

RESULT 1
CSP_PLAFL
ID CSP_PLAFL STANDARD; PRT; 315 AA.
AC P05691;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DE 01-FEB-1994 (Rel. 28, Last annotation update)
DE Circumsporozoite protein (CS) (Fragment).
OS Plasmodium falciparum (isolate 1e5).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5840;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87308186; PubMed=2442154;
RA la Cruz V.F., Lal A.A., McCutchan T.F.;
RT "Sequence variation in putative functional domains of the
RT circumsporozoite protein of Plasmodium falciparum. Implications for
RT vaccine development."
J. Biol. Chem. 262:11935-11939(1987).
RL J. Biol. Chem. 262:11935-11939(1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17802; AAA29538.1; -
KW Malaria; Sporozoite; Repeat.
FT NON_TER 1
FT NON_TER 107 270 40 X 4 AA TANDEM REPEATS OF P-N-A-N.
FT NON_TER 315 315
SQ SEQUENCE 315 AA; 33649 MW; A334DB11FA7FD777 CRC64;

Query Match 100.0%; Score 115; DB 1; Length 315;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNVDPNANPNANPNANP 20
Db 124 NANPNVDPNANPNANPNANP 143

RESULT 2
CSP_PLAFL
ID CSP_PLAFL STANDARD; PRT; 397 AA.
AC P19597; Q25798;
DT 01-FEB-1991 (Rel. 17, Created)

```

DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Circumsporozoite protein precursor (CS).
 OS Plasmodium falciparum (isolate NF54).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NX NCBI_TaxID=5843;
 RA SEQUENCE FROM N.A.
 RX MEDLINE=89345189; PubMed=2668895;
 RA Campbell J.R.;
 RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria
 candidate vaccine antigen."
 RT Nucleic Acids Res. 17:5854-5854(1989).
 RN [2]
 RP REVISIONS.
 RA Campbell J.R.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92155298; PubMed=1346766;
 RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,
 Thomas A.W., Baqar S., Cochran M.A., Thanassi J., Levine M.M.,
 Hackett C.S.;
 RT "Plasmodium falciparum: in vitro characterization and human
 infectivity of a cloned line."
 RT Exp. Parasitol. 74:159-168(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89364986; PubMed=2671723;
 RA Capet P., Gentz R., Matile H., Pink J.R., Stinagaglia F.;
 RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum
 isolate used in malaria vaccine trials."
 RT Mol. Biochem. Parasitol. 35:185-190(1989).
 CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
 SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
 MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
 VERTEBRATE HOST).
 CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
 ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
 WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
 CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X15363; CAA3421.1; -;
 DR EMBL: M83886; AAA29521.1; -;
 DR EMBL: M23982; AAA29527.1; -;
 DR PIR: S05428; S05428.
 DR PIR: A45527; A45527.
 DR InterPro: IPR003067; Circumsporozoite.
 DR InterPro: IPR000884; TSP1.
 DR Pfam: PF00090; TSP_1; 1.
 DR PRINTS: PRO1303; CIRCUMSPOROZOITE.
 DR SMART: SM00209; TSP1; 1.
 KW Malaria; Sporozoite; Repeat; Signal.
 FT SIGNAL 1 16 PROBABLE.
 FT CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.
 FT DOMAIN 101 272 43 X 4 AA TANDEM REPEATS OF N-A-N-P.
 FT CONFLICT 194 194 A -> ANPNANPNA (IN REF. 4).
 SQ SEQUENCE 397 AA; 42646 MM; 9EB1146F59BCEA3 CRC64;
 Query Match 100.0%; Score 115; DB 1; Length 397;
 Best Local Similarity 100.0%; Pred. No. 4.2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NANPVDPNANPNANPNANP 20
 ||||||||||||||||||||

Db 193 NANPVDPNANPNANPNANP 212
 RESULT 3
 CSP_PLAFA STANDARD; PRT; 412 AA.
 ID CSP_PLAFA
 AC P02893;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Circumsporozoite protein precursor (CS).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NX NCBI_TaxID=5843;
 RA SEQUENCE FROM N.A.
 RX MEDLINE=84250215; PubMed=6204383;
 RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
 Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
 Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;
 RT "Structure of the gene encoding the immunodominant surface antigen on
 the sporozoite of the human malaria parasite Plasmodium falciparum."
 RT Science 225:593-599(1984).
 CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
 SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
 MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
 VERTEBRATE HOST).
 CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
 ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
 WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
 CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: K02194; AAA29524.1; -;
 DR PIR: A03388; OZ20AF.
 DR InterPro: IPR003067; Circumsporozoite.
 DR InterPro: IPR000884; TSP1.
 DR Pfam: PF00090; TSP_1; 1.
 DR PRINTS: PRO1303; CIRCUMSPOROZOITE.
 DR SMART: SM00209; TSP1; 1.
 KW Malaria; Sporozoite; Repeat; Signal.
 FT SIGNAL 1 16 PROBABLE.
 FT CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.
 FT DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.
 SQ SEQUENCE 412 AA; 44420 MM; 1EEED3DE9065F8 CRC64;
 Query Match 100.0%; Score 115; DB 1; Length 412;
 Best Local Similarity 100.0%; Pred. No. 4.4e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NANPVDPNANPNANPNANP 20
 ||||||||||||||||||||

Db 204 NANPVDPNANPNANPNANP 223
 RESULT 4
 CSP_PLAFA STANDARD; PRT; 424 AA.
 ID CSP_PLAFA
 AC P13614;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Circumsporozoite protein precursor (CS).
 OS Plasmodium falciparum (isolate t4 / Thailand).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NX NCBI_TaxID=5846;


```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87315205; PubMed=3306373;
RA "del Portillo H.A., Nussenzweig R.S., Enea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from
RL Thailand."
RMol. Biochem. Parasitol. 24:289-294(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M19752; AAA29555.1; -.
CC PIR: A54533; A54533.
CC InterPro: IPR003067; Crcmsprzoite.
CC InterPro: IPR000884; TSP1.
CC Pfam: PF00090; tsp_1; 1.
CC PRINTS: PR01303; CRCMSPRZOITE.
CC SMART: SM00209; TSP1; 1.
CC Malaria; Sporozoite; Repeat; Signal.
CC SIGNAL 16 PROBABLE.
CC CHAIN 17 424 CIRCUMSPOROZOITE PROTEIN.
CC DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
CC SQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;
CC -----
Query Match 100.0%; Score 115; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 4.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NANPNVDNPNANPNANP 20
Db 140 NANPNVDNPNANPNANP 159
-----
RESULT 5
CSP_PLAFW STANDARD; PRT; 442 AA.
P08307;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5846;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87115616; PubMed=3543671;
RA Lockyer M.J., Schwarz R.T.;
RT "Strain variation in the circumsporozoite protein gene of Plasmodium
RT falciparum."
RL Mol. Biochem. Parasitol. 22:101-108(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M15505; AAA29554.1; -.
CC PIR: A54529; A54529.
CC InterPro: IPR003067; Crcmsprzoite.
CC InterPro: IPR000884; TSP1.
CC Pfam: PF00090; tsp_1; 1.
CC PRINTS: PR01303; CRCMSPRZOITE.
CC SMART: SM00209; TSP1; 1.
CC Malaria; Sporozoite; Repeat; Signal.
CC SIGNAL 16
CC CHAIN 17 442 CIRCUMSPOROZOITE PROTEIN.
CC DOMAIN 130 320 47 X 4 AA TANDEM REPEATS OF N-A-N-P.
CC SQUENCE 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;
CC -----
Query Match 100.0%; Score 115; DB 1; Length 442;
Best Local Similarity 100.0%; Pred. No. 4.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NANPNVDNPNANPNANP 20
Db 142 NANPNVDNPNANPNANP 161
-----
RESULT 6
CSP_PLARE STANDARD; PRT; 388 AA.
P26694;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91201303; PubMed=2016283;
RA Lal A.A., Goldman I.F.;
RT "Circumsporozoite protein gene from Plasmodium reichenowi, a
RT chimpanzee malaria parasite evolutionarily related to the human
RT malaria parasite Plasmodium falciparum."
RL J. Biol. Chem. 266:6686-6689(1991).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M60972; AAA29561.1; -.
CC PIR: A39756; A39756.
CC InterPro: IPR003067; Crcmsprzoite.
CC InterPro: IPR000884; TSP1.
CC Pfam: PF00090; tsp_1; 1.
CC PRINTS: PR01303; CRCMSPRZOITE.
CC SMART: SM00209; TSP1; 1.
CC -----

```

KM Malaria; Sporozoite; Repeat; Signal.
 FT SIGNAL 1 16 PROBABLE.
 FT CHAIN 17 388 CIRCUMSPOROZOITE PROTEIN.
 FT DOMAIN 120 267 62 X 4 AA TANDEM REPEATS OF N-A-N-P.
 SO SEQUENCE 388 AA; 42245 MW; C031EEFB2E35604 CRC64;

Query Match 96.5%; Score 111; DB 1; Length 388;
 Best Local Similarity 95.0%; Pred. NO. 1.3e-07;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 NANPVNDPNNANPNANPN 20
 Db 148 NANPVNDPNNANPNANPN 167

RESULT 7

CSP_PLABE STANDARD; PRT; 339 AA.
 AC P06915;
 RT 01-JAN-1988 (Rel. 06, Created)
 RT 01-JAN-1988 (Rel. 06, Last sequence update)
 RT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE CIRCUMSPOROZOITE protein precursor (CS).
 OS Plasmodium berghei.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5821;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87089740; PubMed=242395;
 RA Eichinger D.J., Arnot D.E., Tam J.P., Nussenzweig V., Enea V.;
 RT "Circumsporozoite protein of Plasmodium berghei: gene cloning and
 RT identification of the immunodominant epitopes.";
 RL Mol. Cell. Biol. 6:3965-3972(1986).
 CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
 CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
 CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
 CC VERTEBRATE HOST).
 CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
 CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
 CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
 CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; M14135; AAA29577.1; -;
 DR PIR; A25083; OZ20BK.
 DR InterPro; IPR003067; Crcmsprzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp.1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS50092; TSP1; 1.
 KW Malaria; Sporozoite; Repeat; Signal.
 FT SIGNAL 1 23 PROBABLE.
 FT CHAIN 24 339 CIRCUMSPOROZOITE PROTEIN.
 FT DOMAIN 93 196 13 X 8 AA TANDEM REPEATS.
 FT DOMAIN 206 238 16 X 2 AA TANDEM REPEATS OF P-Q.
 SO SEQUENCE 339 AA; 37138 MW; E8068A6D1D9551B CRC64;

Query Match 49.6%; Score 57; DB 1; Length 339;
 Best Local Similarity 63.2%; Pred. NO. 1;
 Matches 12; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 1 NANPVNDPNNANPNANPN 19
 Db 137 NANDPAPPNANDPAPPNAN 155

RESULT 8
 CSP_PLABA STANDARD; PRT; 347 AA.
 ID CSP_PLABA
 AC P23093;

DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE CIRCUMSPOROZOITE protein precursor (CS).
 OS Plasmodium berghei (strain Anka).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90221834; PubMed=2183186;
 RA Lockyer M.J., Davies C.S., Subhler A., Sinden R.E.;
 RT "Nucleotide sequence of the Plasmodium berghei circumsporozoite
 RT protein gene from the ANKA clone 2.34L.";
 RL Nucleic Acids Res. 18:376-376(1990).

CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
 CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
 CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
 CC VERTEBRATE HOST).
 CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
 CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
 CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
 CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; X17606; CAA35608.1; -;
 DR PIR; S07873; OZ20BK.
 DR InterPro; IPR003067; Crcmsprzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp.1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS50092; TSP1; 1.
 KW Malaria; Sporozoite; Repeat; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 347 CIRCUMSPOROZOITE PROTEIN.
 FT DOMAIN 93 204 13 X 8 AA REPEATS.
 FT DOMAIN 214 247 17 X 2 AA REPEATS OF P-Q.
 SO SEQUENCE 347 AA; 37776 MW; 0EC240EE35681AF8 CRC64;

Query Match 49.6%; Score 57; DB 1; Length 347;
 Best Local Similarity 63.2%; Pred. NO. 1;
 Matches 12; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 1 NANPVNDPNNANPNANPN 19
 Db 145 NANDPAPPNANDPAPPNAN 163

RESULT 9

HAGA_PORGI STANDARD; PRT; 2628 AA.
 ID HAGI_PORGI
 AC Q51845;

DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemagglutinin A precursor.
 GN HAGI.

OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;
 OC Porphyromonadaceae; Porphyromonas.
 OX NCBI_TaxID=837;

RESULT	10
SGK_RAT	
ID	SGK_RAT
AC	Q06226
DT	01-FEB-1995 (Rel. 31, Created)
DT	01-FEB-1995 (Rel. 31, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Serine/threonine-protein kinase Sgk (EC 2.7.1.-)
DE	(Serum/glucocorticoid-regulated kinase).
GN	SGK
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Fischer 344;
RC	MEDLINE=93204949; PubMed=8455596;
RA	Webster M.K., Goya L., Ge Y., Malvar A.C., Firestone G.L.;
RT	"Characterization of sgk, a novel member of the serine/threonine
RT	protein kinase gene family, which is transcriptionally induced by
RT	glucocorticoids and serum."
RL	Mol. Cell. Biol. 13:2031-2040(1993).
CC	-1 FUNCTION: PROBABLE PROTEIN KINASE THAT MAY ACT IN RENAL EPITHELIA
CC	TO ACTIVATE APICAL SODIUM CHANNELS (BY SIMILARITY).
CC	-1 TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES WITH HIGHEST
CC	LEVELS IN THE OVARY, THYMUS AND LUNG.
CC	-1 INDUCTION: BY DEXAMETHASONE AND SERUM.
CC	-1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -

RESULT	11
SGK_HUMAN	
ID	SGK_HUMAN
AC	000141; Q8UN56;
DC	15-DEC-1998 (Rel. 37, Created)
DT	15-DEC-1998 (Rel. 37, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Serine/threonine-protein kinase SGK (EC 2.7.1.-)
DE	(Serum/glucocorticoid-regulated kinase).
GN	SGK.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID:9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE-97272242; PubMed-9114008;
RT	Waldegger S., Barth P., Raber G., Lang F.;
RT	"Cloning and characterization of a putative human serine/threonine
RT	protein kinase transcriptionally modified during anisotonic and
RT	isotonic alterations of cell volume.";
RL	Proc. Natl. Acad. Sci. U.S.A. 94:4440-4445(1997).
RN	[2]
RP	SEQUENCE FROM N.A.
RX	MEDLINE-98390195; PubMed-9722955;
RT	Waldegger S., Edel M., Neigl U.O., Barth P., Raber G., Steuer S.;
RT	Utermann G., Palmichl M., Lang F.;
RT	"Genomic organization and chromosomal localization of the human SGK
RT	protein kinase gene.";
RL	Genomics 51:299-302(1998).
RN	[3]
RP	SEQUENCE FROM N.A.
RA	Kim M.K., Kim Y.H., Seo J.M., Lee H.M., Chung H.J., Sohn M.Y.,
RA	Hwang S.Y., Im S.U., Jung E.J., Kim J.C.;
RT	"A catalogue of genes in the human dermal papilla cells as identified
RT	by expressed sequence tags.";
RL	Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
RN	[4]

```
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RA Strauberg R.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLE PROTEIN KINASE THAT MAY ACT IN RENAL EPITHELIA
CC TO ACTIVATE APICAL SODIUM CHANNELS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y10032; CAA71138.1; -
DR EMBL: AJ000512; CAA04146.1; -
DR EMBL: AF153609; AAD41091.1; -
DR EMBL: BC001263; AAH01263.1; -
DR HSSP: P00517; 1YDR.
DR GENE: HGNC:10810; SGK.
DR MIM: 602958; -
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR000961; Pkinase_C.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; Pkinase_1.
DR Pfam: PF00433; Pkinase_C; 1.
DR PRODOM: PD000001; Euk_Pkinase; 1.
DR SMART: SM00133; S_TK_X; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR TRANSFERASE: Serine/threonine-protein kinase; ATP-binding.
DR DOMAIN: 98 355
DR NP_BIND: 104 112 ATP (BY SIMILARITY).
DR BINDING: 127 127 ATP (BY SIMILARITY).
DR ACT_SITE: 222 222 BY SIMILARITY.
DR CONFLICT: 381 381 E -> D (IN REF. 3 AND 4).
SO SEQUENCE 431 AA; 48956 MW; F3697DA5707399D CRC64;

Query Match 47.0%; Score 54; DB 1; Length 431;
Best Local Similarity 42.3%; Pred. No. 3.1;
Matches 11; Conservative 4; Mismatches 5; Indels 6; Gaps 1;

OY 1 NANPVPD-----NANPNANPNANP 20
   ||||:| 1:11:11:11
   70 NANPSPSPSQOINLGPSNPNAKP 95

RESULT 12
SGK_MOUSE
ID SGK_MOUSE STANDARD; PRT; 431 AA.
AC Q9WVC6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine-protein kinase Sgk (EC 2.7.1.1 -)
DE (Serum/glucocorticoid-regulated kinase).
GN SGK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RX MEDLINE=99287894; PubMed=10358046;
RA Naray-Fejes-Toth A., Canessa C., Cleaveland E.S., Aldrich G.,
RA Fejes-Toth G.;
RT "sgk is an aldosterone-induced kinase in the renal collecting duct.
RT Effects on epithelial Na+ channels."
RT J. Biol. Chem. 274:16973-16978(1999).
```

```
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215285; PubMed=10751222;
RA Shigaev A., Asher C., Latier H., Garty H., Reuveny E.;
RT "Regulation of sgk by aldosterone and its effects on the epithelial
RT Na(+) channel."
RL Am. J. Physiol. 278:F613-F619(2000).
CC -1- FUNCTION: PROBABLE PROTEIN KINASE THAT MAY ACT IN RENAL EPITHELIA
CC TO ACTIVATE APICAL SODIUM CHANNELS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF139638; AAD43302.1; -
DR EMBL: AF205855; AAF19429.1; -
DR HSSP: P00517; 1YDR.
DR MGD: MGI:1340062; Sgk.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR000961; Pkinase_C.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; Pkinase_1.
DR Pfam: PF00433; Pkinase_C; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_Pkinase; 1.
DR SMART: SM00133; S_TK_X; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR TRANSFERASE: Serine/threonine-protein kinase; ATP-binding.
DR DOMAIN: 98 355
DR NP_BIND: 104 112 ATP (BY SIMILARITY).
DR BINDING: 127 127 ATP (BY SIMILARITY).
DR ACT_SITE: 222 222 BY SIMILARITY.
SO SEQUENCE 431 AA; 48928 MW; 6DF5B846A4C2734 CRC64;

Query Match 47.0%; Score 54; DB 1; Length 431;
Best Local Similarity 42.3%; Pred. No. 3.1;
Matches 11; Conservative 4; Mismatches 5; Indels 6; Gaps 1;

OY 1 NANPVPD-----NANPNANPNANP 20
   ||||:| 1:11:11:11
   Db 70 NANPSPSPSQOINLGPSNPNAKP 95

RESULT 13
SGK_RABIT
ID SGK_RABIT STANDARD; PRT; 431 AA.
AC Q9XT18;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine-protein kinase Sgk (EC 2.7.1.1 -)
DE (Serum/glucocorticoid-regulated kinase).
GN SGK.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_TaxID=9986;
RX MEDLINE=99287894; PubMed=10358046;
RA Naray-Fejes-Toth A., Canessa C., Cleaveland E.S., Aldrich G.,
RA Fejes-Toth G.;
RT "sgk is an aldosterone-induced kinase in the renal collecting duct.
```

RT Effects on epithelial Na⁺ channels.";
 RL J. Biol. Chem. 274:16973-16978(1999).
 CC -!- FUNCTION: PROBABLE PROTEIN KINASE THAT MAY ACT IN RENAL EPITHELIA
 CC TO ACTIVATE APICAL SODIUM CHANNELS.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: A0139639; AAD43303.1; -.
 DR HSSP: P00517.1YDR.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR000961; Pkinase_C.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR Pfam: PF00433; pkinase.C.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR PRODOM: PD000001; Euk_pkinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00133; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PSS0108; PROTEIN_KINASE_ST; 1.
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 98 355
 FT NP_BIND 104 112 ATP (BY SIMILARITY).
 FT BINDING 127 127 ATP (BY SIMILARITY).
 FT ACT_SITE 222 222 BY SIMILARITY.
 SQ SEQUENCE 431 AA; 48999 MW; 354898A77EEB38FD CRC64;
 Query Match 47.0%; Score 54; DB 1; Length 411;
 Best Local Similarity 42.3%; Pred. No. 3.1;
 Matches 11; Conservative 4; Mismatches 5; Indels 6; Gaps 1;
 QY 1 NANPNDP-----NANPNPNANP 20
 Db 70 NANSPPPSPSQINLGPSSNPHAKP 95
 RESULT 14
 ID P60_LISMO STANDARD; PRT; 484 AA.
 CD P21171; 003493;
 DT 01-NOV-1991 (Rel. 18, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein p60 precursor (Invasion-associated protein).
 GN IAP OR IMAO0582.
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-49.
 RC STRAIN=EGD / Serovar 1/2a;
 RX MEDLINE=90256283; Pubmed=2111287;
 RA Koehler S., Leimweister-Waechter M., Chakraborty T., Lottspeich F.,
 RA Goebel W.;
 RT "The gene coding for protein p60 of Listeria monocytogenes and its
 RT use as a specific probe for Listeria monocytogenes.";
 RL Infect. Immun. 58:1943-1950(1990).
 RN [2]
 RP SEQUENCE FROM N.A., AND DISCUSSION OF SEQUENCE.
 RC STRAIN=Mackness / Serovar 1/2a;
 RX MEDLINE=93094153; Pubmed=1459966;
 RA Budert A., Kuhn W., Goebel W., Koehler S.;
 RT "Structural and functional properties of the p60 proteins from
 RT different Listeria species.";

RL J. Bacteriol. 174:8166-8171(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EGD-e / Serovar 1/2a;
 RX MEDLINE=21537279; Pubmed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Darvar A., Deboux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurge O.,
 RA Eutlier K.-D., Esbl H., Garcia-del Portillo F., Garrido P.,
 RA Gauthier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kuruprat G.,
 RA Medueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schluter T., Simoes N., Trier A.,
 RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
 RT "Comparative genomics of Listeria species.";
 RL Science 294:849-852(2001).
 CC -!- FUNCTION: THIS MAJOR EXTRACELLULAR PROTEIN MAY BE INVOLVED IN THE
 CC INVASION OF NONPROFESSIONAL PHAGOCYTOTIC CELLS BY LISTERIA.
 CC -!- DOMAIN: LYSM REPEATS ARE THOUGHT TO BE INVOLVED IN PEPTIDOLYCAN
 CC BINDING.
 CC -!- SIMILARITY: BELONGS TO THE E.COLI NLPC / LISTERIA P60 FAMILY.
 CC -!- SIMILARITY: CONTAINS 2 LYSM REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X52268; CAA36509.1; -.
 DR EMBL: M80351; CAA25280.1; -.
 DR EMBL: AL591975; CAC98661.1; ALT_INIT.
 DR PIR: A41487; A41487.
 DR Listlist: LMO00582; -.
 DR MEROPS: C40.0PM;
 DR InterPro: IPR002482; LysM.
 DR InterPro: IPR000064; NLPC_P60.
 DR InterPro: IPR003646; SH3_bac.
 DR Pfam: PF00877; NLPC_P60; 2.
 DR Pfam: PF01476; LysM; 2.
 DR SMART: SM00257; LysM; 2.
 DR SMART: SM00287; SH3b; 1.
 KW Repeat; Signal; Complete proteome.
 FT SIGNAL 1 27
 FT CHAIN 28 484
 FT DOMAIN 311 355
 FT REPEAT 30 72
 FT REPEAT 203 245
 FT VARIANT 94 94
 FT VARIANT 167 167 S -> T (IN STRAIN MACKNESS).
 FT VARIANT 196 196 A -> V (IN STRAIN MACKNESS).
 FT VARIANT 326 331 V -> I (IN STRAIN MACKNESS).
 SQ SEQUENCE 484 AA; 50587 MW; 3CC0F90591E14E0F CRC64;
 Query Match 47.0%; Score 54; DB 1; Length 484;
 Best Local Similarity 52.6%; Pred. No. 3.6;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 1 NANPNDPNNANPNANP 19
 Db 339 NTNTNTNTNTNTNTNAN 357
 RESULT 15
 ID YNH8_YEAST STANDARD; PRT; 407 AA.
 AC P53939;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, last annotation update)
DE Hypothetical 45.9 kDa protein in TPM1-MKS1 intergenic region.
GN YNL078W OR N2337.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SZ28C / FY1679;
RX MEDLINE=96267764; PubMed=8701611;
RA Poehmann R., Philippsen P.;
RT "Sequencing a cosmid clone of Saccharomyces cerevisiae chromosome XIV
RT reveals 12 new open reading frames (ORFs) and an ancient duplication
RT of six ORFs."
RL Yeast 12:391-402(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X86470; CAA60180.1; -
DR EMBL: Z71354; CAA95952.1; -
DR SGD: S0005022; YNL078W.
KW Hypothetical protein.
SQ SEQUENCE 407 AA; 45908 MW; 925C7B6063BCE4F1 CRC64;

Query Match 46.1%; Score 53; DB 1; Length 407;
Best Local Similarity 47.4%; Pred. NO. 4;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 NANPNVDPNANPNANPNAN 19
|:|:|:|:|:|:|:|:|:|
Db 42 NSNSNSTNTNSTNTNSTN 60

Search completed: December 6, 2002, 12:44:01
Job time : 7.55738 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:41:26 : Search time 22.9508 Seconds
(without alignments)
179.555 Million cell updates/sec

Title: US-09-931-325c-3
Perfect score: 115
Sequence: 1 NANPNVDPNANPNANPNANP 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match length	ID	Description
1	115	100.0	383	09GPN1	09GPN1 plasmodium
2	115	100.0	393	09G255	09G255 plasmodium
3	115	100.0	408	025729	025729 plasmodium
4	115	100.0	416	025829	025829 plasmodium
5	115	100.0	420	025838	025838 plasmodium
6	115	100.0	420	025831	025831 plasmodium
7	115	100.0	424	099256	099256 plasmodium
8	115	100.0	424	027425	027425 plasmodium
9	115	100.0	432	025827	025827 plasmodium
10	115	100.0	432	027246	027246 plasmodium
11	115	100.0	436	025828	025828 plasmodium
12	115	100.0	436	027325	027325 plasmodium
13	115	100.0	442	025830	025830 plasmodium
14	115	100.0	452	025834	025834 plasmodium
15	84	73.0	126	09DH26	09dh26 meleagr id h
16	84	73.0	1918	08SSW3	08ssw3 dictyostell

17	77	67.0	314	11	054817	054817 mus musculu
18	77	67.0	382	11	088933	088933 mus musculu
19	77	67.0	387	11	092203	092203 mus musculu
20	75	65.2	378	5	09V807	09V807 drosophila
21	70	60.9	169	12	069490	069490 human herpe
22	70	60.9	169	12	069523	069523 human herpe
23	70	60.9	214	5	09V9Y2	09V9Y2 drosophila
24	70	60.9	502	5	09W474	09W474 drosophila
25	70	60.9	530	5	076918	076918 drosophila
26	70	60.9	1667	2	09R9B7	09R9B7 porphyromon
27	70	60.9	1706	2	051839	051839 porphyromon
28	70	60.9	1706	2	051838	051838 porphyromon
29	68	59.1	1571	16	08X962	08X962 escherichia
30	67	58.3	493	16	09R8P5	09R8P5 mycoplasma
31	64	55.7	1171	5	061732	061732 drosophila
32	64	55.7	1178	5	09W3E2	09W3E2 drosophila
33	64	55.7	1548	5	095P10	095P10 dictyostell
34	63	54.8	387	10	09M9V1	09M9V1 arabidopsis
35	63	54.8	1122	5	08SY41	08SY41 drosophila
36	63	54.8	1658	5	09Y021	09Y021 drosophila
37	63	54.8	2274	5	09YVU0	09YVU0 drosophila
38	62	53.9	356	10	09SAK1	09SAK1 arabidopsis
39	62	53.9	1349	16	08X6C1	08X6C1 escherichia
40	61.5	53.5	1231	5	097140	097140 dictyostell
41	61	53.0	146	10	09FY84	09FY84 arabidopsis
42	61	53.0	284	10	09LJ57	09LJ57 arabidopsis
43	61	53.0	501	16	09JY49	09JY49 neisseria m
44	61	53.0	504	2	087783	087783 neisseria m
45	61	53.0	517	16	09JYD9	09JYD9 neisseria m

ALIGNMENTS

RESULT 1
ID 09GPN1 PRELIMINARY; PRT; 383 AA.
AC 09GPN1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FCC-1/HN;
RX MEDLINE=21367847; PubMed=11474012;
RA Zheng C., Xie P., Chen Y.;
RT "Molecular Cloning and Sequencing of the Circumsporozoite Protein Gene from Plasmodium falciparum Strain FCC-1/HN and Expression of the Gene in Mycobacteria."
RT J. Clin. Microbiol. 39:2911-2915(2001).
RL EMBL: AF315469; AAC37074.1; "-"
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PRO1303; CRCMSPRZOITE.
FT NON_TER 1 1
FT NON_TER 383 383
SQ SEQUENCE 383 AA; 40893 MW; 503C5DFDF61A9E27 CRC64;
Query Match 100.0%; Score 115; DB 5; Length 383;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NANPNVDPNANPNANPNANP 20
Db 177 NANPNVDPNANPNANPNANP 196
RESULT 2

099255 ID 099255 PRELIMINARY: PRT: 393 AA.
AC 099255:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Circumsporozoite protein precursor (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91270295; PubMed=2052038;
RA Lockyer M.J.;
RT "Clonal variation in the Plasmodium falciparum circumsporozoite protein gene."
RL Mol. Biochem. Parasitol. 45:179-181(1991).
DR EMBL; M57498; AAA63421.1; -;
DR InterPro: IPR003067; Circumsporzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT SIGNAL. 1 16 POTENTIAL.
FT CHAIN 17 >393 CIRCUMSPOROZOITE PROTEIN.
FT NON_TER 393 393
SQ SEQUENCE 393 AA; 42263 MW; 45169AE773689037 CRC64;
Query Match 100.0%; Score 115; DB 5; Length 393;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NANPVDPNANPNANPNANP 20
DB 185 NANPVDPNANPNANPNANP 204
RESULT 3
025729 ID 025729 PRELIMINARY: PRT: 408 AA.
AC 025729:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Circumsporozoite protein.
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SANTA LUCIA;
RA Gairi S.H., Lal A.A.;
RL Submitted (FEb-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20969; AAA63153.1; -;
DR InterPro: IPR003067; Circumsporzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
RW Malaria.
SQ SEQUENCE 408 AA; 43871 MW; A545BE517822515C CRC64;
Query Match 100.0%; Score 115; DB 5; Length 408;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NANPVDPNANPNANPNANP 20
DB 204 NANPVDPNANPNANPNANP 223

RESULT 4
025829 ID 025829 PRELIMINARY: PRT: 416 AA.
AC 025829:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAD20;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium falciparum from Thai field isolates."
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83172; AAA29550.1; -;
DR InterPro: IPR003067; Circumsporzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
SQ SEQUENCE 416 AA; 44829 MW; D3EF560B2D368DE9 CRC64;
Query Match 100.0%; Score 115; DB 5; Length 416;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NANPVDPNANPNANPNANP 20
DB 208 NANPVDPNANPNANPNANP 227
RESULT 5
025838 ID 025838 PRELIMINARY: PRT: 420 AA.
AC 025838:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=835B;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium falciparum from Thai field isolates."
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83161; AAA29574.1; -;
DR InterPro: IPR003067; Circumsporzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
SQ SEQUENCE 420 AA; 45155 MW; 3A85B92432C2893C CRC64;
Query Match 100.0%; Score 115; DB 5; Length 420;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NANPVDPNANPNANPNANP 20
DB 220 NANPVDPNANPNANPNANP 239

RESULT 6
ID 025831 PRELIMINARY; PRT: 420 AA.
AC 025831;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RA MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
falciparum from Thai field isolates."
L Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL: M83174; AAA29552.1; -;
DR InterPro: IPR003067; Circspproite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PSS0092; TSP1; 1.
SQ SEQUENCE 420 AA; 45318 MW; F7F70F1C4939DEA7 CRC64;
Query Match 100.0%; Score 115; DB 5; Length 420;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NANPVDPNANPNANPNANP 20
Db 140 NANPVDPNANPNANPNANP 159
|||||
RESULT 7
ID 099256 PRELIMINARY; PRT: 424 AA.
AC 099256;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Circumsporozoite protein precursor (fragment).
GN Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91270295; PubMed=2052038;
RA Lockyer M.J.;
RT "Clonal variation in the Plasmodium falciparum circumsporozoite
protein gene."
L Mol. Biochem. Parasitol. 45:179-181(1991).
DR EMBL: M57499; AAA63422.1; -;
DR InterPro: IPR003067; Circspproite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
KW Signal.
FT SIGNAL 1 16
FT CHAIN 17 >424 CIRCUMSPOROZOITE PROTEIN.
FT NON_TER 424 424
SQ SEQUENCE 424 AA; 45610 MW; BF6D0F06C6648BF CRC64;
Query Match 100.0%; Score 115; DB 5; Length 424;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPVDPNANPNANPNANP 20
Db 148 NANPVDPNANPNANPNANP 167
|||||
RESULT 8
ID 027425 PRELIMINARY; PRT: 424 AA.
AC 027425;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=844;
RA MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
the sporozoite of the human malaria parasite Plasmodium falciparum."
L Science 225:593-599(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=844;
RA la Cruz V.F.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=844;
RA Jongwutives S., Tanabe K., Kanbara H.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL: M83169; AAA29547.1; -;
DR EMBL: M83149; AAA29562.1; -;
DR InterPro: IPR003067; Circspproite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
SQ SEQUENCE 424 AA; 45592 MW; F20CEB60636DB99E CRC64;
Query Match 100.0%; Score 115; DB 5; Length 424;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NANPVDPNANPNANPNANP 20
Db 140 NANPVDPNANPNANPNANP 159
|||||
RESULT 9
ID 025827 PRELIMINARY; PRT: 432 AA.
AC 025827;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=838;
RA MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium

RT falciparum from Thai field isolates.";
 RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
 DR EMBL; M83165; AAA29543.1; -.
 DR InterPro: IPR003067; CrCmsprzoite.
 DR InterPro: IPR000884; TSPL.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PRO1303; CrCmsprzoite.
 DR SMART; SM00209; TSPL; 1.
 SQ SEQUENCE 432 AA; 46385 MW; 2CE8D9A68E11945F CRC64;

Query Match 100.0%; Score 115; DB 5; Length 432;
 Best Local Similarity 100.0%; Pred. No. 3.4e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPVVDNPNANPNANPN 20
 DB 140 NANPVVDNPNANPNANPN 159

RESULT 10
 027246 PRELIMINARY; PRT; 432 AA.

AC 027246;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Circumsporozoite protein.
 GN CSP.
 OS Plasmodium falciparum.
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5633;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-946;
 RX MEDLINE=84250215; PubMed=6204383;
 RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
 Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
 Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
 RT "Structure of the gene encoding the immunodominant surface antigen on
 the sporozoite of the human malaria parasite Plasmodium falciparum.";
 RL Science 225:593-599(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-946;
 RA la Cruz V.F.;
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.

Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL; M83155; AAA29568.1; -.
 DR EMBL; M83170; AAA29548.1; -.
 DR EMBL; M83152; AAA29565.1; -.
 DR EMBL; M83158; AAA29571.1; -.
 DR EMBL; M83166; AAA29544.1; -.
 DR EMBL; M83168; AAA29546.1; -.
 DR InterPro: IPR003067; CrCmsprzoite.
 DR InterPro: IPR000884; TSPL.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PRO1303; CrCmsprzoite.
 DR SMART; SM00209; TSPL; 1.
 SQ SEQUENCE 432 AA; 46414 MW; 8787E6005578873A CRC64;

Query Match 100.0%; Score 115; DB 5; Length 432;
 Best Local Similarity 100.0%; Pred. No. 3.4e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPVVDNPNANPNANPN 20
 DB 148 NANPVVDNPNANPNANPN 167

RESULT 11
 ID 025828
 AC 025828; PRELIMINARY; PRT; 436 AA.

DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Circumsporozoite protein.
 GN CSP.
 OS Plasmodium falciparum.
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5633;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-842;
 RX MEDLINE=95077069; PubMed=7985759;
 RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
 RT "Allelic variation in the circumsporozoite protein of Plasmodium
 falciparum from Thai field isolates.";
 RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
 DR EMBL; M83167; AAA29545.1; -.
 DR InterPro: IPR003067; CrCmsprzoite.
 DR InterPro: IPR000884; TSPL.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PRO1303; CrCmsprzoite.
 DR SMART; SM00209; TSPL; 1.
 DR PROSITE; PS50092; TSPL; 1.
 SQ SEQUENCE 436 AA; 46875 MW; F102683C5C1DC85A CRC64;

Query Match 100.0%; Score 115; DB 5; Length 436;
 Best Local Similarity 100.0%; Pred. No. 3.4e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPVVDNPNANPNANPN 20
 DB 140 NANPVVDNPNANPNANPN 159

RESULT 12
 ID 027325
 AC 027325; PRELIMINARY; PRT; 436 AA.

DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Circumsporozoite protein.
 GN CSP.
 OS Plasmodium falciparum.
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5633;
 RN [1]
 RP SEQUENCE FROM N.A.

Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL; M83155; AAA29568.1; -.
 DR EMBL; M83170; AAA29548.1; -.
 DR EMBL; M83152; AAA29565.1; -.
 DR EMBL; M83158; AAA29571.1; -.
 DR EMBL; M83166; AAA29544.1; -.
 DR EMBL; M83168; AAA29546.1; -.
 DR InterPro: IPR003067; CrCmsprzoite.
 DR InterPro: IPR000884; TSPL.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PRO1303; CrCmsprzoite.
 DR SMART; SM00209; TSPL; 1.
 SQ SEQUENCE 432 AA; 46414 MW; 8787E6005578873A CRC64;

Query Match 100.0%; Score 115; DB 5; Length 432;
 Best Local Similarity 100.0%; Pred. No. 3.4e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPVVDNPNANPNANPN 20
 DB 148 NANPVVDNPNANPNANPN 167

DR InterPro: IPR003067; Crmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; TSP_1; 1.
DR PRINTS: PR01303; CRMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
SQ SEQUENCE 436 AA; 4668 MW; 5B42FF3348B68655 CRC64;

Query Match 100.0%; Score 115; DB 5; Length 436;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNVDPNANPNANPNANP 20
Db 132 NANPNVDPNANPNANPNANP 151

RESULT 13
Q25830 PRELIMINARY; PRT; 442 AA.

ID Q25830
AC Q25830;
DT 01-NOV-1996 (TREMblrel. 01, Created).
DE 01-DEC-2001 (TREMblrel. 19, Last sequence update)
GN Circumsporozoite protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=To/94;
RX MEDLINE=95077069; PubMed=7985759;
RA Tongvutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
falciparum from Thai field isolates."
RL Am. J. Trop. Med. Hyg. 51:659-668 (1994).
DR EMBL: M83173; AAA29551.1; -
DR InterPro: IPR003067; Crmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; TSP_1; 1.
DR PRINTS: PR01303; CRMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
SQ SEQUENCE 442 AA; 47414 MW; BFAF9D939D7862FF CRC64;

Query Match 100.0%; Score 115; DB 5; Length 442;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNVDPNANPNANPNANP 20
Db 142 NANPNVDPNANPNANPNANP 161

RESULT 14

Q25834 PRELIMINARY; PRT; 452 AA.
AC Q25834;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DE 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
GN Circumsporozoite protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=827;
RX MEDLINE=95077069; PubMed=7985759;
RA Tongvutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
falciparum from Thai field isolates."
RL Am. J. Trop. Med. Hyg. 51:659-668 (1994).

DR EMBL: M83156; AAA29569.1; -
DR InterPro: IPR003067; Crmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; TSP_1; 1.
DR PRINTS: PR01303; CRMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
SQ SEQUENCE 452 AA; 48431 MW; 6E739D6C53223805 CRC64;

Query Match 100.0%; Score 115; DB 5; Length 452;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNVDPNANPNANPNANP 20
Db 132 NANPNVDPNANPNANPNANP 151

RESULT 15

Q9DH26 PRELIMINARY; PRT; 126 AA.
AC Q9DH26;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DE 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Hypothetical 14.0 kDa protein.
GN HY076 OR HVT099.
OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=37108;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC126;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.K.;
RT "The genome of turkey herpesvirus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF291866; AAC45826.1; -
DR EMBL: AF291866; AAC45821.1; -
KW Hypothetical protein.
SQ SEQUENCE 126 AA; 13979 MW; 751496E819279A72 CRC64;

Query Match 73.0%; Score 84; DB 12; Length 126;
Best Local Similarity 70.0%; Pred. No. 0.00017;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 NANPNVDPNANPNANPNANP 20
Db 87 NPNNPNPNPNPNPNPNPNP 106

Search completed: December 6, 2002, 12:45:19
Job time : 24.9508 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:38:35 : Search time 30.2951 Seconds
(Without alignments)
92.367 Million cell updates/sec

Title: US-09-931-325C-79
Perfect score: 112
Sequence: 1 EYLKNIKQNSLSTFEMSPCSVT 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	21	AAU93880	P. falciparum univ
2	112	100.0	21	AAU87735	Malarial universal
3	112	100.0	22	AAU87718	Malarial universal
4	112	100.0	26	AAU93980	Modified portion o
5	112	100.0	26	AAU87793	Modified C-terminu
6	112	100.0	171	AAU93972	Immunogenic Hbc ch
7	112	100.0	191	AAU93974	Chimeric Hepatitis
8	112	100.0	195	AAU93975	Chimeric Hepatitis
9	108	96.4	20	AAW29734	Universal malarial
10	108	96.4	20	AAW61554	T-cell epitope 2.

11	108	96.4	20	AAU93860	P. falciparum MSPI
12	108	96.4	20	AAU93924	Hepatitis B virus
13	108	96.4	20	AAU87777	Malarial epitope s
14	108	96.4	40	AAU87213	P. falciparum deriv
15	108	96.4	424	AAU87797	RTS* protein. Syn
16	103	92.0	27	AAU93981	Modified portion o
17	103	92.0	27	AAU87794	Modified C-terminu
18	99	88.4	20	AAU93803	P. falciparum CS B
19	99	88.4	20	AAU87778	Malarial epitope s
20	93	83.0	20	AAU9548	Vaccine related MH
21	93	83.0	54	AAU71644	Circumsporozoite R
22	93	83.0	180	AAU07290	Circumsporozoite a
23	93	83.0	184	AAU07289	Circumsporozoite a
24	93	83.0	250	AAU60412	Synthetic peptide
25	93	83.0	309	AAU3175	NS1_81-RLfauth.
26	93	83.0	319	AAU07945	NS1_81-RLfauth. plas
27	93	83.0	319	AAU3176	NS1_81-RLfauth. p
28	93	83.0	327	AAU3177	NS1_81-RLfauth + (
29	93	83.0	335	AAU3178	NS1_81(NANP)ARLfa
30	93	83.0	335	AAU3179	NS1_81(NANP)ARLfa
31	93	83.0	411	AAU83144	Sequence encoded b
32	93	83.0	412	AAU60416	CS protein of mala
33	93	83.0	412	AAU80835	Sequence encoded b
34	93	83.0	424	AAU37796	RTS protein. Syn
35	82	73.2	402	AAU70709	Plasmodium cynomol
36	81	72.3	21	AAU93887	P. vivax universal
37	81	72.3	21	AAU87742	Malarial universal
38	79	70.5	161	AAU07287	Circumsporozoite a
39	77	68.8	160	AAU07288	Circumsporozoite a
40	77	68.8	378	AAU30609	Plasmodium vivax c
41	77	68.8	429	AAU90064	Antigenic protein
42	77	68.8	1807	AAU85697	Recombinant protei
43	77	68.8	2028	AAU85698	Recombinant protei
44	75	67.0	378	AAU70708	Sequence encoding
45	72	64.3	19	AAU93861	Plasmodium vivax T

ALIGNMENTS

RESULT 1	AAU93880	standard; Peptide: 21 AA.
ID	AAU93880	
AC	AAU93880;	
DT	02-JUL-2002	(first entry)
XX		
DE	P. falciparum universal T cell epitope Pf-UTC.	
XX		
KW	Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; Hbc;	
KM	vaccine; B cell epitope; T cell epitope; immunostimulant.	
XX		
OS	Plasmodium falciparum.	
PN	WO200214478-A2.	
XX		
PD	21-FEB-2002.	
XX		
PF	16-AUG-2001; 2001WO-US41759.	
XX		
PR	16-AUG-2000; 2000US-225843P.	
PR	22-AUG-2000; 2000US-226867P.	
PR	15-AUG-2001; 2001US-0950915.	
PA	(APOV-) APOVIA INC.	
XX		
PI	Birkett AJ;	
XX		
DR	WPI; 2002-257601/30.	
XX		
PT	Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric hepatitis B core protein, displays immunogenic epitopes at N-terminus,	

PT Hbc immunogenic loop with linker for conjugated epitope and C-terminus
PT
XX
PS
XX
Example 1; Page 110; 289pp; English.

CC The invention relates to a recombinant hepadnavirus nucleocapsid protein,
CC i.e. a chimeric hepatitis B core (Hbc) protein (1), displaying one or
CC more immunogenic epitopes at the N-terminus, Hbc immunogenic loop (L) or
CC C-terminalus, or having a heterologous linker for a conjugated epitope in
CC (L), and containing a Cys residue at, or near, the C-terminus that
CC confers enhanced stability to the particles. A vaccine comprising (1) is
CC useful for inducing an immune response in an inoculated host animal, by
CC inoculating a host animal with the vaccine, and maintaining that
CC inoculated animal for a time period sufficient for that animal to
CC develop an immune response. The immunogenic particles formed using (1)
CC are substantially free of binding to nucleic acids, and are most stable
CC than the particle formed from otherwise identical Hbc chimera that lacks
CC the C-terminal residue or in which a C-terminal Cys is replaced by
CC another residue. The chimera particles are most stable on storage in
CC aqueous compositions that are particles of similar sequence that lack any
CC C-terminal Cys residues. The chimera molecule exhibits the self-assembly
CC not exhibiting the nucleic acid binding of those native particles, and
CC excellent B cell and T cell immunogenicities. The chimera particles are
CC typically prepared in higher yield than similar particles that are free
CC of a C-terminal Cys. The particles are often far more immunogenic than
CC the similar conjugates that lack a C-terminal Cys. Immunogenicities of
CC particles assembled from the chimera molecules are enhanced as compared to
CC similar particles assembled from chimera molecules lacking at least one
CC C-terminal Cys. AAU93802-AAU93997 represent immunogenic Hbc particles
CC amino acid sequences and related sequences of the invention.
XX
SQ Sequence 21 AA:

Query Match 100.0%; Score 112; DB 23; Length 21;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IEYLNKIONSLSTEWSPCSVT 21
|
Db 1 IEYLNKIONSLSTEWSPCSVT 21

RESULT 2
AAU87735
ID AAU87735 standard; Peptide; 21 AA.
XX
AC AAU87735;
XX
21-MAY-2002 (first entry)

Malarial universal T cell epitope #4.

XX Hepatitis B virus; nucleocapsid protein; Hbc; hepatitis B virus core;
XX B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoacide;
XX circumsporozoite; human immunodeficiency virus type 1; human; squirrel;
XX woodchuck.

XX Plasmodium falciparum.

XX WO200213765-A2.

XX 21-FEB-2002.

XX 16-AUG-2001; 2001WO-US25625.

XX 16-AUG-2000; 2000US-225813P.

XX 15-AUG-2001; 2001US-0931325.

XX (APOV-) APOVIA INC.

XX Birkett AJ;

XX WPT; 2002-241832/29.

DR N-PSDB; ABEK44248.

XX Recombinant hepatitis B virus core protein chimera molecule, useful to
PT induce antibodies to malarial parasites, comprises malaria-specific
PT T-cell epitope and is engineered for enhanced stability -
PT

XX Example 1; Fig 2B; 197pp; English.

CC The invention relates to a recombinant hepatitis B virus core (Hbc)
CC protein chimera molecule that contains 4 peptide-linked amino acid residue
CC sequence domains. The molecule of the invention contains a region
CC constituting a B cell epitope of the circumsporozoite protein of a
CC species of the parasite, Plasmodium. The chimera sequence is useful as an
CC immunogen for inducing antibodies to the malaria-causing parasite,
CC Plasmodium, particularly P. falciparum and P. vivax. Sequences
CC AAU87695-AAU87804 represent peptide epitopes of the invention.
XX

SQ Sequence 21 AA:

Query Match 100.0%; Score 112; DB 23; Length 21;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IEYLNKIONSLSTEWSPCSVT 21
|
Db 1 IEYLNKIONSLSTEWSPCSVT 21

RESULT 3

AAU87718
ID AAU87718 standard; Peptide; 22 AA.

XX AAU87718;

XX 21-MAY-2002 (first entry)

XX Malarial universal T cell epitope #1.

XX Hepatitis B virus; nucleocapsid protein; Hbc; hepatitis B virus core;
XX B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoacide;
XX circumsporozoite; human immunodeficiency virus type 1; human; squirrel;
XX woodchuck.

XX Plasmodium falciparum.

XX WO200213765-A2.

XX 21-FEB-2002.

XX 16-AUG-2001; 2001WO-US25625.

XX 16-AUG-2000; 2000US-225813P.

XX 15-AUG-2001; 2001US-0931325.

XX (APOV-) APOVIA INC.

XX Birkett AJ;

XX WPT; 2002-241832/29.

XX Recombinant hepatitis B virus core protein chimera molecule, useful to
PT induce antibodies to malarial parasites, comprises malaria-specific
PT T-cell epitope and is engineered for enhanced stability -
PT

XX Claim 51; Page 38; 197pp; English.

CC The invention relates to a recombinant hepatitis B virus core (Hbc)
CC protein chimera molecule that contains 4 peptide-linked amino acid residue
CC sequence domains. The molecule of the invention contains a region
CC constituting a B cell epitope of the circumsporozoite protein of a
CC species of the parasite, Plasmodium. The chimera sequence is useful as an
CC immunogen for inducing antibodies to the malaria-causing parasite,
CC Plasmodium, particularly P. falciparum and P. vivax. Sequences

CC AAU87695-AAU87804 represent peptide epitopes of the invention.
XX
SQ Sequence 22 AA;
Query Match 100.0%; Score 112; DB 23; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 IEYLNKIONSLSTEMSPCSVT 21
DB 2 IEYLNKIONSLSTEMSPCSVT 22
RESULT 4
AAU93980
ID AAU93980 standard; Peptide: 26 AA.
AC AAU93980;
XX
XX 02-JUL-2002 (first entry)
DE Modified portion of Hepatitis B core.
XX
XX Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; HBC;
KM vaccine; B cell epitope; T cell epitope; immunostimulant.
XX
XX Hepatitis B virus.
OS WO200214478-A2.
XX
XX 21-FEB-2002.
PD
PE 16-AUG-2001; 2001WO-US41759.
XX
XX 16-AUG-2000; 2000US-225843P.
PR 22-AUG-2000; 2000US-226867P.
PR 15-AUG-2001; 2001US-0930915.
XX
XX (APOV-) APOVIA INC.
PA
XX
XX Birkett AJ;
PI
XX
XX WPI: 2002-257601/30.
DR
XX
XX Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric
PT hepatitis B core protein, displays immunogenic epitopes at N-terminus,
PT HBC immunogenic loop with linker for conjugated epitope and C-terminus
PT
PS
XX
XX Example 22; Page 163; 289pp; English.
XX
XX The invention relates to a recombinant hepadnavirus nucleocapsid protein,
CC i.e. a chimeric hepatitis B core (HBC) protein (1), displaying one or
CC more immunogenic epitopes at the N-terminus, HBC immunogenic loop (L) or
CC C-terminus, or having a heterologous linker for a conjugated epitope in
CC (L), and containing a Cys residue at, or near, the C-terminus that
CC confers enhanced stability to the particles. A vaccine comprising (1) is
CC useful for inducing an immune response in an inoculated host animal, by
CC inoculating a host animal with the vaccine, and maintaining that
CC inoculated animal for a time period sufficient for that animal to
CC develop an immune response. The immunogenic particles formed using (1)
CC are substantially free of binding to nucleic acids, and are most stable
CC than the particle formed from otherwise identical Hbc chimera that lacks
CC the C-terminal residue or in which a C-terminal Cys is replaced by
CC another residue. The chimera particles are most stable on storage in
CC aqueous compositions that are particles of similar sequence that lack any
CC C-terminal Cys residues. The chimera molecule exhibits the self-assembly
CC not exhibiting the nucleic acid binding of those native particles, and
CC excellent B cell and T cell immunogenicities. The chimera particles are
CC typically prepared in higher yield than similar particles that are free
CC of a C-terminal Cys. The particles are often far more immunogenic than
CC the similar conjugates that lack a C-terminal Cys. Immunogenicities of
CC particles assembled from the chimera molecules are enhanced as compared to

CC similar particles assembled from chimera molecules lacking at least one
CC C-terminal Cys. AAU93802-AAU93997 represent immunogenic HBC particles
CC amino acid sequences and related sequences of the invention.
XX
SQ Sequence 26 AA;
Query Match 100.0%; Score 112; DB 23; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 IEYLNKIONSLSTEMSPCSVT 21
DB 6 IEYLNKIONSLSTEMSPCSVT 26
RESULT 5
AAU87793
ID AAU87793 standard; Peptide: 26 AA.
AC AAU87793;
XX
XX 21-MAY-2002 (first entry)
DE Modified C-terminus of Hepatitis B virus nucleocapsid protein #13.
XX
XX Hepatitis B virus; nucleocapsid protein; HBC; hepatitis B virus core;
KM B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoacide;
KM circumsporozoite; human immunodeficiency virus type I; human; squirrel;
KM woodchuck.
XX
XX Hepatitis B virus.
OS Synthetic.
XX
XX WO200213765-A2.
XX
XX 21-FEB-2002.
PD
PE 16-AUG-2001; 2001WO-US25625.
XX
XX 16-AUG-2000; 2000US-225813P.
PR 15-AUG-2001; 2001US-0931325.
XX
XX (APOV-) APOVIA INC.
PA
XX
XX Birkett AJ;
PI
XX
XX WPI: 2002-241832/29.
DR
XX
XX Recombinant hepatitis B virus core protein chimera molecule, useful to
PT induce antibodies to malarial parasites, comprises malaria-specific
PT T-cell epitope and is engineered for enhanced stability -
PT
PS
XX
XX Example 13; Page 128; 197pp; English.
XX
XX The invention relates to a recombinant hepatitis B virus core (HBC)
CC protein chimera molecule that contains a peptide-linked amino acid residue
CC sequence domain. The molecule of the invention contains a region
CC constituting a B cell epitope of the circumsporozoite protein of a
CC species of the parasite, Plasmodium. The chimera sequence is useful as an
CC immunogen for inducing antibodies to the malaria-causing parasite,
CC Plasmodium, particularly P. falciparum and P. vivax. Sequences
CC AAU87695-AAU87804 represent peptide epitopes of the invention.
XX
XX
SQ Sequence 26 AA;
Query Match 100.0%; Score 112; DB 23; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 IEYLNKIONSLSTEMSPCSVT 21
DB 6 IEYLNKIONSLSTEMSPCSVT 26

```

XX RESULT 6
XX ID AU093972
XX AU093972 standard; Peptide: 171 AA.
XX
XX AU093972:
XX
XX 02-JUL-2002 (first entry)
XX
XX Immunogenic Hbc chimeric particle #9.
XX
XX Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; Hbc;
XX vaccine; B cell epitope; T cell epitope; immunostimulant.
XX
XX Plasmodium falciparum.
XX
XX WO200214478-A2.
XX
XX 21-FEB-2002.
XX
XX 16-AUG-2001; 2001WO-USA1759.
XX
XX 16-AUG-2000; 2000US-225843P.
XX
XX 22-AUG-2000; 2000US-226867P.
XX
XX 15-AUG-2001; 2001US-0930915.
XX
XX (APOV-) APOVIA INC.
XX
XX Birkett AJ;
XX
XX WPI; 2002-257601/30.
XX
XX Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric
XX hepatitis B core protein, displays immunogenic epitopes at N-terminus,
XX Hbc immunogenic loop with linker for conjugated epitope and C-terminus
XX
XX
XX Example 4; Page 273; 289pp; English.
XX
XX The invention relates to a recombinant hepadnavirus nucleocapsid protein,
XX 1.e. a chimeric hepatitis B core (Hbc) protein (I), displaying one or
XX more immunogenic epitopes at the N-terminus, Hbc immunogenic loop (L) or
XX C-terminus, or having a heterologous linker for a conjugated epitope in
XX (L), and containing a Cys residue at, or near, the C-terminus that
XX confers enhanced stability to the particles. A vaccine comprising (I) is
XX useful for inducing an immune response in an inoculated host animal, by
XX inoculating a host animal with the vaccine, and maintaining that
XX inoculated animal for a time period sufficient for that animal to
XX develop an immune response. The immunogenic particles formed using (I)
XX are substantially free of binding to nucleic acids, and are most stable
XX than the particle formed from otherwise identical Hbc chimera that lacks
XX the C-terminal residue or in which a C-terminal Cys is replaced by
XX another residue. The chimera particles are most stable on storage in
XX aqueous compositions that are particles of similar sequence that lack any
XX C-terminal Cys residues. The chimera molecule exhibits the self-assembly
XX not exhibiting the nucleic acid binding of those native particles, and
XX excellent B cell and T cell immunogenicities. The chimera particles are
XX typically prepared in higher yield than similar particles that are free
XX of a C-terminal Cys. The particles are often far more immunogenic than
XX the similar conjugates that lack a C-terminal Cys. Immunogenicities of
XX particles assembled from the chimera molecules are enhanced as compared to
XX similar particles assembled from chimera molecules lacking at least one
XX C-terminal Cys. AU093802-AU093997 represent immunogenic Hbc particles
XX amino acid sequences and related sequences of the invention.
XX
XX Sequence 171 AA;
XX
XX Query Match 100.0%; Score 112; DB 23; Length 171;
XX Best Local Similarity 100.0%; Pred. No. 3,3e-09;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX 1 IEYLNKIONSLSTEWSPGCVT 21
XX
XX ||||||||||||||||||||

```

Query Match	Best Local Similarity	Score	DB	Length
Matches 21; Conservative 0;	100.0%;	112.0;	DB 23;	191;
	100.0%;	112.0;	Pred. NO. 3.8e-05;	
	0;	Indels 0;	Gaps 0;	

DT 19-OCT-1998 (first entry)
XX
DE T-cell epitope 2.
XX
KM T-cell; malaria; immunogenic; anti-malarial; prophylactic immunity.
XX
OS Plasmodium falciparum
XX
PN WO9831382-A1.
XX
PD 23-JUL-1998.
XX
PE 21-JAN-1998; 98WO-US01527.
XX
PR 21-JAN-1997; 97US-0033916.
XX
PA (UYNY) UNIV NEW YORK STATE.
XX
PI Moreno A, Nardin E;
WPI; 1998-413810/35.
PT New immunogenic compositions for malaria - comprise malaria derived
PT peptide comprising universal T-cell epitope which elicits
XX
XX anti-malarial T-cell response
XX
PS Claim 5; Page 25; 38pp: English.
XX
CC The T-cell epitope derived from malaria can be used in an immunogenic
CC composition. The T-cell epitope elicits an anti-malarial T-cell response
CC in mammals of diverse genetic backgrounds. The composition can be used
CC as a vaccine to confer prophylactic or therapeutic immunity against
CC malaria. They may also be used to inhibit the propagation of a malarial
CC organism in a susceptible animal.
XX
SQ Sequence 20 AA:
Query Match 96.4%; Score 108; DB 19; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 EYLKNIQSLSTWSPSCSVT 21
Db 1 EYLKNIQSLSTWSPSCSVT 20
RESULT 11
AAU93860 standard; Peptide: 20 AA.
AAU93860;
DT 02-JUL-2002 (first entry)
XX
DE P. falciparum MSP1 T cell epitope #4.
XX
KM Immunogenic; hepatitis B core; HBC;
KM vaccine; B cell epitope; T cell epitope; immunostimulant.
XX
OS Plasmodium falciparum.
XX
PN WO200214478-A2.
XX
PD 21-FEB-2002.
XX
PE 16-AUG-2001; 2001WO-US41759.
XX
PR 16-AUG-2000; 2000US-225843P.
PR 22-AUG-2000; 2000US-226867P.
PR 15-AUG-2001; 2001US-0930915.
XX
PA (APOV-) APOVIA INC.
XX

PI Birkett AJ;
XX
DR WPI; 2002-257601/30.
XX
PT Novel recombinant hepatitis B core protein, displaying one or
PT hepatitis B core protein, displays immunogenic epitopes at N-terminus,
PT HBC immunogenic loop with linker for conjugated epitope and C-terminus
XX
XX
PS Disclosure; Page 43; 289pp: English.
XX
CC The invention relates to a recombinant hepatitis B core protein, displaying one or
CC i.e. a chimeric hepatitis B core (HBC) protein (1), displaying one or
CC more immunogenic epitopes at the N-terminus, HBC immunogenic loop (L) or
CC C-terminus, or having a heterologous linker for a conjugated epitope in
CC (L), and containing a Cys residue at, or near, the C-terminus that
CC confers enhanced stability to the particles. A vaccine comprising (1) is
CC useful for inducing an immune response in an inoculated host animal, by
CC inoculating a host animal with the vaccine, and maintaining that
CC inoculated animal for a time period sufficient for that animal to
CC develop an immune response. The immunogenic particles formed using (1)
CC are substantially free of binding to nucleic acids, and are most stable
CC than the particle formed from otherwise identical HBC chimera that lacks
CC the C-terminal residue or in which a C-terminal Cys is replaced by
CC another residue. The chimera particles are most stable on storage in
CC aqueous compositions that are particles of similar sequence that lack any
CC C-terminal Cys residues. The chimera molecule exhibits the self-assembly
CC not exhibiting the nucleic acid binding of those native particles, and
CC excellent B cell and T cell immunogenicities. The chimera particles are
CC typically prepared in higher yield than similar particles that are free
CC of a C-terminal Cys. The particles are often far more immunogenic than
CC the similar conjugates that lack a C-terminal Cys. Immunogenicities of
CC particles assembled from the chimera molecules are enhanced as compared to
CC similar particles assembled from chimera molecules lacking at least one
CC C-terminal Cys. AAU93802-AAU93997 represent immunogenic HBC particles
CC amino acid sequences and related sequences of the invention.
XX
SQ Sequence 20 AA:
Query Match 96.4%; Score 108; DB 23; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 EYLKNIQSLSTWSPSCSVT 21
Db 1 EYLKNIQSLSTWSPSCSVT 20
RESULT 12
AAU93924 standard; Peptide: 20 AA.
AAU93924;
DT 02-JUL-2002 (first entry)
XX
DE Hepatitis B virus HBC149 peptide #14.
XX
KM Immunogenic; hepatitis B core; HBC;
KM vaccine; B cell epitope; T cell epitope; immunostimulant.
XX
OS Hepatitis B virus.
XX
PN WO200214478-A2.
XX
PD 21-FEB-2002.
XX
PE 16-AUG-2001; 2001WO-US41759.
XX
PR 16-AUG-2000; 2000US-225843P.
PR 22-AUG-2000; 2000US-226867P.
PR 15-AUG-2001; 2001US-0930915.
XX

(APOV-) APOVIA INC.

Birkett AJ:

WPI: 2002-257601/30.

Novel recombinant hepatitis B core protein, displays immunogenic epitopes at N-terminus, HBC immunogenic loop with linker for conjugated epitope and C-terminus

Example 8; Page 130; 289pp; English.

The invention relates to a recombinant hepatitis B core (HBC) protein (I), displaying one or more immunogenic epitopes at the N-terminus, HBC immunogenic loop (U) or C-terminus, or having a heterologous linker for a conjugated epitope in (U), and containing a Cys residue at, or near, the C-terminus that confers enhanced stability to the particles. A vaccine comprising (I) is useful for inducing an immune response in an inoculated host animal, by inoculating a host animal with the vaccine, and maintaining that inoculated animal for a time period sufficient for that animal to develop an immune response. The immunogenic particles formed using (I) are substantially free of binding to nucleic acids, and are most stable than the particle formed from otherwise identical HBC chimera that lacks the C-terminal residue or in which a C-terminal Cys is replaced by another residue. The chimera particles are most stable on storage in aqueous compositions that are particles of similar sequence that lack any C-terminal Cys residues. The chimera molecule exhibits the self-assembly not exhibiting the nucleic acid binding of those native particles, and excellent B cell and T cell immunogenicities. The chimera particles are typically prepared in higher yield than similar particles that are free of a C-terminal Cys. The particles are often far more immunogenic than the similar conjugates that lack a C-terminal Cys. Immunogenicities of particles assembled from the chimera molecules are enhanced as compared to similar particles assembled from chimera molecules lacking at least one C-terminal Cys. AAU93802-AAU9397 represent Immunogenic HBC particles amino acid sequences and related sequences of the invention.

Sequence 20 AA:

Query Match 96.4%; Score 108; DB 23; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0

2 EYLKNTKNSLSTEWSPCSVT 21
|||||
1 EYLKNTKNSLSTEWSPCSVT 20

RESULT 13
AAU87777
ID AU87777 standard; Peptide: 20 AA.
XX AU87777:
DT 21-MAY-2002 (first entry)
XX
DE Malarial epitope sequence #4.
XX
KW Hepatitis B virus: nucleocapsid protein: HBC; hepatitis B virus core:
KM B cell epitope: T cell epitope; malaria; HIV-1; Plasmodium; protozoasidae;
KM circumsporozoite; human immunodeficiency virus type I; human; squirrel;
woodchuck.
XX Plasmodium falciparum.
OS
PN MO200213765-A2.
PD
XX 21-FEB-2002.
PF 16-AUG-2001; 2001WO-US25625.
XX

PR	16-AUG-2000; 2000US-225813P.
PR	15-AUG-2001; 2001US-093132S.
XX	
PA	(ABOV-) AFOVIA INC.
XX	
PI	Birkett AJ;
XX	
DR	WPI: 2002-241832/29.
XX	
PT	Recombinant hepatitis B virus core protein chimera molecule, useful to induce antibodies to malarial parasites, comprises malaria-specific T-cell epitope and is engineered for enhanced stability -
PT	
PS	Claim 25; Page 110; 197pp; English.
XX	
CC	The invention relates to a recombinant hepatitis B virus core (Hbc)
CC	protein chimera molecule that contains 4 peptide-linked amino acid residue
CC	sequence domains. The molecule of the invention contains a region
CC	constituting a B cell epitope of the circumsporozoite protein of a
CC	species of the parasite, Plasmodium. The chimera sequence is useful as an
CC	immunogen for inducing antibodies to the malaria-causing parasite,
CC	Plasmodium, particularly P. falciparum and P. vivax. Sequences
CC	AU087695-AAU87804 represent peptide epitopes of the invention.
XX	
SO	Sequence 20 AA;
	Query Match 96.4%; Score 108; DB 23; Length 20;
	Best Local Similarity 100.0%; Pred. No. 1.2e-09;
	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	2 EYLKTIQNSLSTEWSPCSVT 21 EYLKTIQNSLSTEWSPCSVT 20
DB	1 EYLKTIQNSLSTEWSPCSVT 20
RESULT 14	
ID	AA087213
XX	AA087213 standard; peptide; 40 AA.
AC	
XX	AA087213;
DT	16-MAY-1996 (first entry)
DE	P.falciparum derived presented KEKE-like motif contg. peptide.
KX	Proteasome; activation; cell-mediated immunity; immunogen; tolerance;
KW	KEKE motif; interleukin.
OS	Plasmodium falciparum.
XX	
XX	Key Location/Qualifiers
FT	Region 1..20
FT	/note= "KEKE-like sequence"
FT	Misc-difference 21
FT	/note= "undefined linker sequence of 34 amino acids"
FT	Peptide 22..40
FT	/note= "presented peptide"
XX	
PN	WO9527058-A1.
PD	12-OCT-1995.
PF	01-APR-1994; 94WO-US03591.
PR	01-APR-1994; 94WO-US03591.
XX	
PA	(UTAH) UNIV UTAH.
XX	
PI	Realini CA, Rechsteiner MC;
XX	
DR	WPI: 1995-358633/46.

PT	DNA encoding human proteasome activator - used to elicit
PT	cell-mediated immunity or tolerance to a selected immunogenic peptide
XX	
PS	Disclosure; Page 45; 71pp; English.
XX	
CC	AAR87210-R87215 are, presented peptides contg. a KEKE-like motif.
CC	KEKE motif contg. peptides can be used in a method for inducing cell-
CC	-mediated immunity against or tolerance to specific epitopes using
CC	plasmids encoding a human proteasome activator (PA) and appropriate
CC	epitope-bearing peptides adjacent to presentation marker peptides
CC	(lysine and glutamine rich peptides termed KEK motifs peptides, that
CC	mark adjacent peptides for presentation). The method can be used to
CC	elicit cell- mediated immunity or tolerance to pathogen-encoded peptides
CC	or tumour specific antigens.
XX	
SQ	Sequence 40 AA;
	Query Match 96.4%; Score 108; DB 16; Length 40;
	Best Local Similarity 100.0%; Pred. No. 2,6e-09;
	Matches 20; Conservative .0; Mismatches 0; Indels 0; Gaps 0;
OY	2 EYLXNKIIONSLSTEWSPCSVT 21 EYLXNKIIONSLSTEWSPCSVT 20
Db	1 EYLXNKIIONSLSTEWSPCSVT 20
RESULT 15	
AAR37797	
ID	AAR37797 standard; Protein; 424 AA.
XX	
AC	AAR37797;
XX	
DT	27-SEP-1993 (first entry)
XX	
DE	RTS* protein.
XX	
KW	RTS; expression cassette; hybrid protein; S. cerevisiae; TDH3;
KW	cloning; circumsporozoite protein; CSP; Plasmodium falciparum;
KW	strain 768; hepatitis B virus; HBV; adw serotype; pres2 protein;
KW	S protein.
OS	Synthetic.
XX	
FT	Key Location/Qualifiers
FT	1
FT	/note= "Derived from S. cerevisiae TDH3 gene sequence"
FT	Region 2..4
FT	/note= "Cloning artefact"
FT	Protein 5..193
FT	/note= "Represents amino acids 210-398 of the CSP of P. falciparum"
FT	Region 194..197
FT	/note= "Carboxy terminal amino acids from HBV (adw serotype) pres2 protein"
FT	Protein 198..424
FT	/note= "S protein of HBV (adw serotype)"
PN	MO310152-A.
PD	
PD	27-MAY-1993.
PF	11-NOV-1992; 92WO-EP02591.
PR	16-NOV-1991; 91GB-0024390.
PR	27-FEB-1992; 92US-0842694.
PA	(SMIK) SMITHLINE BEECAM BIOLOGICALS.
PI	Cohen J, De Wilde M;
WI	WPI, 1993-182494/22.
N-PSDB:	AAO42567.

xx		Hybrid protein comprising Plasmodium circumsporozoite protein and HbsAg - useful as a vaccine for treating patients susceptible to Plasmodium infections
pt		
pn		
xx		
ps		Disclosure: Fig 9; 59pp; English.
xx		
cc		This sequence represents the RTS ⁺ hybrid protein which is encoded by the RTS ⁺ expression cassette. This hybrid consists of a methionine residue derived from S. cerevisiae PH3 gene sequence, three amino acids, Met-Ala-Pro, derived from a nucleotide sequence created by the cloning procedure used to construct the hybrid gene, a stretch of 189 amino acids representing amino acids 210 to 398 of the circumsporozoite protein (CSP) of Plasmodium falciparum strain NF54, an amino acid A19 created by the cloning procedure, four amino acids, Pro-Val-Thr-Ileu, representing the four carboxy terminal residues of hepatitis B virus (HBV), adw serotype, prest protein, and a stretch of 226 amino acids specifying the S protein of HBV, adw serotype.
cc		This protein, and RTS (see also AAR37196), may be combined with an adjuvant and used in a vaccine for preventing plasmodium infections.
cc		The vaccines produce a humoral response and also a cellular immune response.
cc		
xx		
sq	Sequence	424 AA;
qy	Query Match	96.4%; Score 108; DB 14; Length 424;
	Best Local Similarity	100.0%; Pred. No. 3.8e-08;
	Matches 20; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
db	2 EYLNKIONSISTEMSPCSV	21
	EYLNKIONSISTEMSPCSV	143

Search completed: December 6, 2002, 12:43:35
Job time : 30.2951 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:42:06 ; Search time 12.0492 Seconds
(Without alignments)
51.280 Million cell updates/sec

Title: US-09-931-325C-79
Perfect score: 112
Sequence: 1 EYLNKIONSLSTEWSPCSVT 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCrUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	96.4	424	2	US-08-760-797A-3
2	108	96.4	424	4	US-08-932-929B-3
3	93	83.0	412	1	US-08-313-288B-18
4	93	83.0	423	2	US-08-760-797A-1
5	93	83.0	424	4	US-08-932-929B-1
6	77	68.8	20	1	US-07-848-636B-3
7	77	68.8	378	6	5171843-9
8	62	55.4	478	3	US-08-155-88B-2
9	50	44.6	46	1	US-08-395-602A-5
10	50	44.6	46	2	US-08-021-625D-5
11	50	44.6	525	4	US-09-369-602A-21
12	49	43.8	18	1	US-08-395-602A-1
13	49	43.8	18	2	US-08-021-625D-1
14	49	43.8	23	1	US-08-395-602A-2
15	49	43.8	23	2	US-08-021-625D-2
16	49	43.8	23	4	US-08-986-659B-9
17	49	43.8	23	4	US-08-986-659B-33
18	49	43.8	33	4	US-08-986-659B-34
19	49	43.8	114	1	US-08-309-604-2
20	49	43.8	126	1	US-08-395-602A-4
21	49	43.8	126	2	US-08-021-625D-4
22	49	43.8	559	1	US-08-313-288B-14
23	46	41.1	38	1	US-08-444-005-18
24	46	41.1	38	3	US-08-985-526-30
25	46	41.1	459	1	US-08-313-288B-15
26	46	41.1	656	1	US-08-444-005-15
27	46	41.1	656	4	US-09-069-023-28

28	45	40.2	232	6	5171843-7	Patent No. 5171843
29	44	39.3	9	1	US-07-646-531D-1	Sequence 1, Appl
30	44	39.3	9	1	US-07-646-531D-8	Sequence 8, Appl
31	44	39.3	9	1	US-07-646-531D-15	Sequence 15, Appl
32	44	39.3	9	1	US-07-646-531D-16	Sequence 16, Appl
33	44	39.3	9	2	US-08-488-273-1	Sequence 8, Appl
34	44	39.3	9	2	US-08-488-273-8	Sequence 13, Appl
35	44	39.3	9	2	US-08-488-273-13	Sequence 14, Appl
36	44	39.3	9	2	US-08-488-273-14	Sequence 1, Appl
37	44	39.3	9	2	US-08-858-971-1	Sequence 2, Appl
38	44	39.3	9	4	US-09-197-770B-2	Sequence 14, Appl
39	44	39.3	9	4	US-09-197-770B-14	Sequence 14, Appl
40	44	39.3	9	6	5426100-1	Patent No. 5426100
41	44	39.3	9	6	5426100-8	Patent No. 5426100
42	44	39.3	227	4	US-09-182-145-15	Sequence 15, Appl
43	44	39.3	228	4	US-09-182-145-19	Sequence 19, Appl
44	44	39.3	228	4	US-09-182-145-77	Sequence 77, Appl
45	44	39.3	229	4	US-09-182-145-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1
US-08-760-797A-3
Sequence 3, Application US/08760797A
Patent No. 5928902
GENERAL INFORMATION:
APPLICANT: De Wilde, Michel
TITLE OF INVENTION: Hybrid Protein Between CS
TITLE OF INVENTION: from Plasmidum and HBSAG
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smltline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fastseq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,797A
FILING DATE: 04-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/442,612
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumelster, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-760-797A-3
Query Match 96.4%; Score 108; DB 2; Length 424;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 2 EYLNKIONSLSTEWSPCSVT 21

Db 124 EYLNKIONSLSTWSPSCSVT 143

RESULT 2

US-08-932-929B-3
Sequence 3, Application US/08932929B
Patent No. 6169171

GENERAL INFORMATION:

APPLICANT: De Wilde, Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
TITLE OF INVENTION: From Plasmodium and HBSAG
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,929B
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/760,797
FILING DATE: 04-DEC-1996
APPLICATION NUMBER: 08/442,612
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1FWC2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-932-929B-3

Query Match 96.4%; Score 108; DB 4; Length 424;

Best Local Similarity 100.0%; Pred. No. 2.6e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EYLNKIONSLSTWSPSCSVT 21

Db 124 EYLNKIONSLSTWSPSCSVT 143

RESULT 3

US-08-313-288B-18
Sequence 18, Application US/08313288B
Patent No. 5750502

GENERAL INFORMATION:

APPLICANT: Jessell, Thomas M. and Avihu Klar
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: USA
ZIP: 10036
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526

STATE: New York
COUNTRY: USA
ZIP: 10036
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-313-288B-18

Query Match 83.0%; Score 93; DB 1; Length 412;

Best Local Similarity 80.0%; Pred. No. 4.1e-06;

Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EYLNKIONSLSTWSPSCSVT 21

Db 333 QYLNKIONSLSTWSPSCSVT 352

RESULT 4

US-08-760-797A-1
Sequence 1, Application US/08760797A
Patent No. 5928902

GENERAL INFORMATION:

APPLICANT: De Wilde, Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
TITLE OF INVENTION: From Plasmodium and HBSAG
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA

US-08-760-797A-1

Query Match 96.4%; Score 108; DB 4; Length 424;

Best Local Similarity 100.0%; Pred. No. 2.6e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EYLNKIONSLSTWSPSCSVT 21

Db 333 QYLNKIONSLSTWSPSCSVT 352

RESULT 4

US-08-760-797A-1
Sequence 1, Application US/08760797A
Patent No. 5928902

GENERAL INFORMATION:

APPLICANT: De Wilde, Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
TITLE OF INVENTION: From Plasmodium and HBSAG
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA

APPLICANT: De Wilde, Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
TITLE OF INVENTION: From Plasmodium and HBSAG
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA

STATE: New York
COUNTRY: USA
ZIP: 10036
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526

TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-760-797A-1

Query Match 83.0%; Score 93; DB 2; Length 423;
Best Local Similarity 80.0%; Pred. No. 4.2e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 EYLKIKNSLSTWSPCSVT 21
DB 127 QYLKIKNSLSTWSPCSVT 146

RESULT 5

S-08-932-929B-1
Sequence 1, Application US/08932929B
Patent No. 6169171
GENERAL INFORMATION:
APPLICANT: De Wilde, Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
TITLE OF INVENTION: From Plasmidium and HBSAG
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,929B
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/760,797
FILING DATE: 04-DEC-1996
APPLICATION NUMBER: 08/442,612
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumelster, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1FWC2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-932-929B-1

Query Match 83.0%; Score 93; DB 4; Length 424;
Best Local Similarity 80.0%; Pred. No. 4.2e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
OY 2 EYLKIKNSLSTWSPCSVT 21
DB 128 QYLKIKNSLSTWSPCSVT 147

RESULT 6

US-07-848-636B-3
Sequence 3, Application US/07848636B
Patent No. 5599543
GENERAL INFORMATION:
APPLICANT: Hoffman, Stephen L
APPLICANT: Charoenvit, Yupin
APPLICANT: Jones, Trevor R
TITLE OF INVENTION: A PHARMACEUTICAL COMPOSITION CONTAINING
TITLE OF INVENTION: FOUR AMINO ACID EPIPOPE PROTECTIVE AGAINST PLASMODIUM
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Naval Medical Research & Development Command
STREET: 8901 Wisconsin Ave Bldg 1, T-12.
CITY: Bethesda
STATE: Maryland
COUNTRY: USA
ZIP: 20889-5606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/848,636B
FILING DATE: 09-MAR-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,551
FILING DATE: 06-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Spevack, A. David
REGISTRATION NUMBER: 24,743
REFERENCE/DOCKET NUMBER: N.C. 72,634
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 295-6759
TELEFAX: (301) 295-1022
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
PUBLICATION INFORMATION:
DOCUMENT NUMBER: US 5,095,093
FILING DATE: 10-MAR-1992
US-07-848-636B-3

Query Match 68.8%; Score 77; DB 1; Length 20;
Best Local Similarity 60.0%; Pred. No. 3.4e-05;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
OY 2 EYLKIKNSLSTWSPCSVT 21
DB 1 EYLDKVRATVGTWTPCSVT 20

RESULT 7
5171843-9
Patent No. 5171843
APPLICANT: NUSSENZWEIG, VICTOR
TITLE OF INVENTION: IMMUNOGENIC POLYPEPTIDE AND METHOD FOR
PURIFYING IT
NUMBER OF SEQUENCES: 13
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/175,112

;; FILING DATE: 30-MAR-1988
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 754,645
;; FILING DATE: 9-JUL-1985
;; APPLICATION NUMBER: 115,634
;; FILING DATE: 26-OCT-1987
;; APPLICATION NUMBER: 649,903
;; FILING DATE: 12-SEP-1984
;; SEQ ID NO: 9
;; LENGTH: 378
5171843-9

Query Match 68.8%; Score 77; DB 6; Length 378;
Best Local Similarity 60.0%; Pred. No. 0.00086;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 2 EYLKQNSLSTWSPCSVT 21
DB 300 EYLKQNSLSTWSPCSVT 319

SUFT 8
US-08-155-888-2

;; Sequence 2, Application US/08155888
;; Patent No. 6066623
;; GENERAL INFORMATION:
;; APPLICANT: Hoffman, Stephen L.
;; APPLICANT: Hedstrom, Richard C.
;; APPLICANT: Sedegrah, Martha
;; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE PROTECTIVE
;; TITLE OF INVENTION: AGAINST MALARIA, METHODS OF PROTECTION AND VECTOR FOR
;; TITLE OF INVENTION: DELIVERING POLYNUCLEOTIDE VACCINES
;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Naval Medical Res. & Dev. Cmd.
;; STREET: Bldg 1, T-12 8901 Wisconsin Ave.
;; CITY: Bethesda
;; STATE: Maryland
;; COUNTRY: USA
;; ZIP: 20889-5606
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/155,888
;; FILING DATE:
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Spevack, A. David
;; REGISTRATION NUMBER: 24,743
;; REFERENCE/DOCKET NUMBER: N.C. 75,851
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 295-6759
;; TELEFAX: (202) 295-1022
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 478 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-155-888-2

Query Match 55.4%; Score 62; DB 3; Length 478;
Best Local Similarity 47.6%; Pred. No. 0.18;
Matches 10; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EYLKQNSLSTWSPCSVT 21
DB 400 LEFKQNSLSTWSPCSVT 420

RESULT 9
US-08-395-602A-5
;; Sequence 5, Application US/08395602A
;; Patent No. 5766899
;; GENERAL INFORMATION:
;; APPLICANT: Kuo, M. Tien
;; APPLICANT: Ding, Zhi-Ming
;; TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into
;; TITLE OF INVENTION: Liver Cells
;; NUMBER OF SEQUENCES: 5
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Arnold, White & Durkee
;; STREET: P.O. Box 4433
;; CITY: Houston
;; STATE: Texas
;; COUNTRY: United States of America
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,602A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-395-602A-5

Query Match 44.6%; Score 50; DB 1; Length 46;
Best Local Similarity 53.3%; Pred. No. 0.82;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 IONSLSSTWSPCSVT 21
DB 22 VDPNANPEWSPCSVT 36

RESULT 10
US-08-021-625D-5
;; Sequence 5, Application US/08021625D
;; Patent No. 5976851
;; GENERAL INFORMATION:
;; APPLICANT: Kuo, M. Tien
;; APPLICANT: Ding, Zhi-Ming
;; TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into
;; TITLE OF INVENTION: Liver Cells
;; NUMBER OF SEQUENCES: 5
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Arnold, White & Durkee
;; STREET: P.O. Box 4433
;; CITY: Houston
;; STATE: Texas
;; COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30


```

? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/021,625D
? FILING DATE: 16-FEB-1993
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Highlander, Steven L.
? REGISTRATION NUMBER: 37,642
? REFERENCE/DOCKET NUMBER: UTSIC:410/HYL
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (512) 418-3000
? TELEFAX: (713) 789-2679
? TELEX: 79-0924
? INFORMATION FOR SEQ ID NO.: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 46 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? SUBMITTER: US-08-021-625D-5
Oy      7 IONS1STEMSPCSV 21          44.6%, Score 50; DB 2; Length 46;
       : : : |||||         Best Local Similarity   53.3%;
Db      22 VDPNANPEWSPCSV 36          Pred. No. 0.82;
                                         Matches    8; Conservative    3; Mismatches    4; Indels    0; Gaps    0;

```

```

RESULT 11
US-09-369-364A-21
: Sequence 21, Application US/09369364A
: Patent No. 6391610
: GENERAL INFORMATION:
: APPLICANT: Apce, Suneel
: APPLICANT: Hurskainen, Tiina L.
: APPLICANT: Hirohata, Satoshi
: TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
: FILE REFERENCE: 26473/4007/10-30-00
: CURRENT APPLICATION NUMBER: US/09/369,364A
: CURRENT FILING DATE: 1999-08-06
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 21
: LENGTH: 525
: TYPE: PR1
: ORGANISM: Homo sapiens ADAMTS-R1
S-09-369-364A-21

Query Match          44.6%; Score 50; DB 4; Length 525;
Best Local Similarity 72.7%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0.

Cy      11 LSTWSPCSVT 21
      1:|||||:|
Db      441 LAQWSPCTVT 451

RESULT 12
US-08-395-602A-1
: Sequence 1, Application US/08395602A
: Patent No. 5766899
: GENERAL INFORMATION:
: APPLICANT: Kuo, M. Tien
: APPLICANT: Ding, Zhi-Ming
: TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: P.O. Box 4433
: CITY: Houston
: STATE: Texas

```

COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,602A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-395-602A-1

```

Query Match          43.8%; Score 49; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 0 41;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      14  EWSPCSVT 21
        ,  |||||
Db      1  EWSPCSVT 8

RESULT 13
US-08-021-625D-1
: Sequence 1, Application US/08021625D
: Patent No. 5976851
: GENERAL INFORMATION:
  APPLICANT: Kuo, M. Tien
  TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into
  TITLE OF INVENTION: Liver Cells
  NUMBER OF SEQUENCES: 5
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Arnold, White & Durkee
  STREET: P.O. Box 4433
  CITY: Houston
  STATE: Texas
  COUNTRY: United States of America
  ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/021.625D
  FILING DATE: 16-FEB-1993
  CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Highlander, Steven L.
    REGISTRATION NUMBER: 37,642
  REFERENCE/DOCKET NUMBER: DTSC:410/HYL
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (512) 418-3000
    TELEFAX: (713) 789-2679
    TELEX: 79-0924
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 18 amino acids

```

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-021-625D-1

Query Match 43.8%; Score 49; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 EMSPCSVT 21
|||||
DB 1 EMSPCSVT 8

RESULT 14
US-08-395-602A-2
Sequence 2, Application US/08395602A
Patent No. 5766899
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into
TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,602A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-395-602A-2

Query Match 43.8%; Score 49; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 EMSPCSVT 21
|||||
DB 1 EMSPCSVT 8

RESULT 15
US-08-021-625D-2
Sequence 2, Application US/08021625D
Patent No. 5976851
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into

TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,625D
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-021-625D-2

Query Match 43.8%; Score 49; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 EMSPCSVT 21
|||||
DB 1 EMSPCSVT 8

Search completed: December 6, 2002, 12:46:44
Job time : 13.0492 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:44:06 ; Search time 6.54098 Seconds
(without alignments)
52.146 Million cell updates/sec

Title: US-09-931-325C-79
Perfect score: 112
Sequence: 1 IEXLINKIONSLESTFSPCSVT 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/PTCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	44.6	53	10	US-09-800-729-164
2	50	44.6	54	10	US-09-858-081-6
3	50	44.6	54	10	US-09-858-068-6
4	50	44.6	54	12	US-10-014-070-9
5	50	44.6	525	9	US-09-978-295A-301
6	50	44.6	525	9	US-09-978-697-301
7	50	44.6	525	9	US-09-978-192A-301
8	50	44.6	525	10	US-09-918-171A-21
9	50	44.6	525	12	US-10-052-586-436
10	50	44.6	1745	10	US-09-800-729-89
11	50	44.6	1762	9	US-10-044-807-2
12	50	44.6	1766	10	US-09-764-853-554
13	50	44.6	63	12	US-10-036-869-30
14	50	44.6	656	10	US-09-862-027-30
15	50	44.6	9	12	US-10-042-202-11
16	50	44.6	59	10	US-09-864-761-40588
17	50	44.6	465	10	US-09-869-515-12
18	50	44.6	476	10	US-09-869-515-6
19	50	44.6	969	10	US-09-969-515-10

20 45 40.2 980 10 US-09-969-515-4
21 45 40.2 1213 10 US-09-969-515-8
22 45 40.2 1224 10 US-09-969-515-2
23 45 40.2 2150 10 US-09-321-987B-2
24 45 40.2 2155 10 US-09-800-729-135
25 44 39.3 58 9 US-10-010-408-11
26 44 39.3 226 10 US-09-945-676-8
27 44 39.3 227 9 US-10-010-408-13
28 44 39.3 250 9 US-10-010-408-2
29 44 39.3 250 10 US-09-915-582-53
30 44 39.3 250 10 US-09-915-582-69
31 44 39.3 285 9 US-09-992-598A-272
32 44 39.3 285 9 US-09-989-293A-272
33 44 39.3 285 10 US-09-989-722-272
34 44 39.3 285 10 US-09-989-723-272
35 44 39.3 285 10 US-09-989-729-272
36 44 39.3 285 10 US-09-989-727-272
37 44 39.3 285 10 US-09-989-731-272
38 44 39.3 285 10 US-09-989-732-272
39 44 39.3 285 10 US-09-991-073-272
40 44 39.3 285 10 US-09-990-442-272
41 44 39.3 285 10 US-09-991-163-272
42 44 39.3 285 10 US-09-993-604-272
43 44 39.3 285 10 US-09-990-456-272
44 44 39.3 285 10 US-09-989-721-272
45 44 39.3 1588 9 US-10-000-512-2

ALIGNMENTS

RESULT 1
US-09-800-729-164
Sequence 164, Application US/09800729
Patent No. US20020068319A1
GENERAL INFORMATION:
APPLICANT: NI et al.
TITLE OF INVENTION: 32 Human secreted proteins
FILE REFERENCE: P2044PI
CURRENT APPLICATION NUMBER: US/09/800, 729
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: PCT/US00/26013
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/155, 709
PRIOR FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 217
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 164
LENGTH: 53
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-729-164
Query Match 44.6%; Score 50; DB 10; Length 53;
Best Local Similarity 72.7%; Pred. No. 0.33;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 11 LSTEMSPCSVT 21
Db 2 LAQEWSPCTVT 12
RESULT 2
US-09-858-081-6
Sequence 6, Application US/09858081
Patent No. US20020072490A1
GENERAL INFORMATION:
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Cook, William James
APPLICANT: Siles-Santiago, Inmaculada
TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEINASE
FILE REFERENCE: 10448-049001

```

: CURRENT APPLICATION NUMBER: US/09/858,081
: CURRENT FILING DATE: 2001-05-15
: PRIOR APPLICATION NUMBER: 60/204,159
: PRIOR FILING DATE: 2000-05-15
: PRIOR APPLICATION NUMBER: 60/204,160
: PRIOR FILING DATE: 2000-05-15
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 54
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: consensus sequence
US-09-858-081-6

Query Match          44.6%; Score 50; DB 10; Length 54;
Best Local Similarity 88.9%; Pred. No. 0.34;
Matches      8; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

      13 TEMSPCSVY 21
      :|||||
      Db           4 SEMSPCSVY 12

RESULT 3
US-09-858-068-6
: Sequence 6, Application US/09858068
: Patent No. US20020076778A1
: GENERAL INFORMATION:
: APPLICANT: Kapeller-Liebermann, Rosana
: APPLICANT: Cook, William James
: TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEASE
: TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
: FILE REFERENCE: 10448-057001
: CURRENT APPLICATION NUMBER: US/09/858,068
: CURRENT FILING DATE: 2001-05-15
: PRIOR APPLICATION NUMBER: 60/204,159
: PRIOR FILING DATE: 2000-05-15
: PRIOR APPLICATION NUMBER: 60/204,160
: PRIOR FILING DATE: 2000-05-15
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 54
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: consensus sequence
US-09-858-068-6

Query Match          44.6%; Score 50; DB 10; Length 54;
Best Local Similarity 88.9%; Pred. No. 0.34;
Matches      8; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

      13 TEMSPCSVY 21
      :|||||
      Db           4 SEMSPCSVY 12

RESULT 4
US-10-014-070-9
: Sequence 9, Application US/10014070
: Patent No. US20020119555A1
: GENERAL INFORMATION:
: APPLICANT: Bandaru, Rajasehkar
: APPLICANT: Curtis, Rory A.J.
: APPLICANT: Spurling, Heidi Lynn
: APPLICANT: Millennium Pharmaceuticals, Inc.
: TITLE OF INVENTION: 53014, A Human Metalloprotease Family
: TITLE OF INVENTION: Member and Uses Therefor
: FILE REFERENCE: MPI2000-523PIKCP1(M)
```

```

: CURRENT APPLICATION NUMBER: US/10/014,070
: CURRENT FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: 60/258,373
: PRIOR FILING DATE: 2000-12-22
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 54
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: consensus
US-10-014-070-9

Query Match          44.6%; Score 50; DB 12; Length 54;
Best Local Similarity 88.9%; Pred. No. 0.34;
Matches      8; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

      13 TEMSPCSVY 21
      :|||||
      Db           4 SEMSPCSVY 12

RESULT 5
US-09-978-295A-301
: Sequence 301, Application US/09978295A
: Patent No. US20020156006A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Guiney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James
: APPLICANT: Paonli, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2630P1C11
: CURRENT APPLICATION NUMBER: US/09/978,295A
: CURRENT FILING DATE: 2001-10-15
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
```

PRIOR APPLICATION	NUMBER: 60/0825258
PRIOR FILING DATE:	1998-04-21
PRIOR APPLICATION	NUMBER: 60/0825659
PRIOR FILING DATE:	1998-04-21
PRIOR APPLICATION	NUMBER: 60/0827074
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION	NUMBER: 60/0828004
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION	NUMBER: 60/0827000
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION	NUMBER: 60/0827977
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION	NUMBER: 60/0833222
PRIOR FILING DATE:	1998-04-28
PRIOR APPLICATION	NUMBER: 60/0833592
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION	NUMBER: 60/0834959
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION	NUMBER: 60/0834566
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION	NUMBER: 60/0834939
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION	NUMBER: 60/0835454
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION	NUMBER: 60/0835554
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION	NUMBER: 60/0835586
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION	NUMBER: 60/0835599
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION	NUMBER: 60/0835599
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION	NUMBER: 60/0835000
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION	NUMBER: 60/0837422
PRIOR FILING DATE:	1998-04-30
PRIOR APPLICATION	NUMBER: 60/0843666
PRIOR FILING DATE:	1998-05-06
PRIOR APPLICATION	NUMBER: 60/0846373
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION	NUMBER: 60/0846393
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION	NUMBER: 60/0846400
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION	NUMBER: 60/0846427
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION	NUMBER: 60/0846433
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION	NUMBER: 60/0853333
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION	NUMBER: 60/0853389
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION	NUMBER: 60/0853222
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION	NUMBER: 60/0855848
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION	NUMBER: 60/0857900
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION	NUMBER: 60/0856888
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION	NUMBER: 60/0857979
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION	NUMBER: 60/0855800
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION	NUMBER: 60/0855800
PRIOR FILING DATE:	1998-05-15

; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
Query Match 44.6%; Score 50; DB 9; Length 525;
Best Local Similarity 72.7%; Pred. No. 3.9;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Oy 11 LSTEMSPCSVT 21
Db 441 LAOWSPCTVT 451
RESULT 6
US-09-978-697-301
; Sequence 301, Application US/09978697
; Patent No. US20020169284A1
GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT APPLICATION NUMBER: US/09/978, 697
PRIOR FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804

;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-5-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 44.6%; Score 50; DB 9; Length 525;
Best Local Similarity 72.7%; Pred. No. 3.9;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 11 LSTEMSPCSVT 21
Db 441 LAQEMSPCTVT 451
RESULT 7
US-09-978-192A-301
; Sequence 301, Application US/09978192A
; Patent No. US2002017753A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Bolstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, Paul J.
; APPLICANT: Gunney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C9
; CURRENT APPLICATION NUMBER: US/09/978,192A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 2001-07-30
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23

PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/083666
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 44.6%; Score 50; DB 9; Length 525;
Best Local Similarity 72.7%; Pred. No. 3.9;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 11 LSTEMSPCSVT 21
Db 441 LAQWSPCTVT 451


```
RESULT 8
US-09-918-171A-21
: Sequence 21, Application US/09918171A
: Patent No. US20020110894A1
: GENERAL INFORMATION:
: APPLICANT: Apte, Suneel
: APPLICANT: Hurskainen, Tiina L.
: APPLICANT: Hirohata, Satoshi
: TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
: FILE REFERENCE: 26473/04193
: CURRENT APPLICATION NUMBER: US/09/918.171A
: CURRENT FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 09/369,364
: PRIOR FILING DATE: 1999-08-06
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 21
: LENGTH: 525
: TYPE: PRT
: ORGANISM: Homo sapiens ADAMTS-R1
US-09-918-171A-21

Query Match          44.6%; Score 50; DB 10; Length 525;
Best Local Similarity 72.7%; Pred. No. 3.9;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0:

Qy 11 LSTWSPCSVT 21
: : : : : : : :
Db 441 LAQWSPCTVT 451

RESULT 9
US-10-052-586-436
: Sequence 436, Application US/10052586
: Patent No. US20020127584A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Tian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3430R1C1
: CURRENT FILING DATE: 2002-01-15
: PRIOR APPLICATION NUMBER: 60/059263
: PRIOR FILING DATE: 1997-09-18
: PRIOR APPLICATION NUMBER: 60/059266
: PRIOR FILING DATE: 1997-09-18
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/063120
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063121
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063486
: PRIOR FILING DATE: 1997-10-21
: PRIOR APPLICATION NUMBER: 60/063540
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063541
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063544
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063564
```

```
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063734
: PRIOR FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: 60/063870
: PRIOR FILING DATE: 1997-10-31
: PRIOR APPLICATION NUMBER: 60/064103
: PRIOR FILING DATE: 1997-10-31
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066120
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/066466
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 60/066772
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 60/069335
: PRIOR FILING DATE: 1997-12-11
: PRIOR APPLICATION NUMBER: 60/069425
: PRIOR FILING DATE: 1997-12-12
: PRIOR APPLICATION NUMBER: 60/069870
: PRIOR FILING DATE: 1997-12-17
: PRIOR APPLICATION NUMBER: 60/068017
: PRIOR FILING DATE: 1997-12-18
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/078866
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078939
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079664
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079786
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/080107
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080194
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080327
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/080333
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/081049
: PRIOR FILING DATE: 1998-04-08
: PRIOR APPLICATION NUMBER: 60/081070
: PRIOR FILING DATE: 1998-04-08
: PRIOR APPLICATION NUMBER: 60/081195
: PRIOR FILING DATE: 1998-04-09
: PRIOR APPLICATION NUMBER: 60/081838
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: 60/082568
: PRIOR FILING DATE: 1998-04-21
: PRIOR APPLICATION NUMBER: 60/082569
: PRIOR FILING DATE: 1998-04-21
: PRIOR APPLICATION NUMBER: 60/082704
: PRIOR FILING DATE: 1998-04-22
: PRIOR APPLICATION NUMBER: 60/082797
: PRIOR FILING DATE: 1998-04-22
: PRIOR APPLICATION NUMBER: 60/083322
: PRIOR FILING DATE: 1998-04-28
: PRIOR APPLICATION NUMBER: 60/083495
: PRIOR FILING DATE: 1998-04-29
: PRIOR APPLICATION NUMBER: 60/083496
: PRIOR FILING DATE: 1998-04-29
: PRIOR APPLICATION NUMBER: 60/083499
: PRIOR FILING DATE: 1998-04-29
: PRIOR APPLICATION NUMBER: 60/083559
: PRIOR FILING DATE: 1998-04-29
: PRIOR APPLICATION NUMBER: 60/084366
: PRIOR FILING DATE: 1998-05-05
```

PRIOR APPLICATION NUMBER: 60/084414
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084639
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084640
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084643
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085573
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085580
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085582
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085700
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/086023
 PRIOR FILING DATE: 1998-05-18
 PRIOR APPLICATION NUMBER: 60/086392
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/086486
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/087098
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087208
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087609
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087759
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087827
 PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088655
 PRIOR FILING DATE: 1998-06-09
 PRIOR APPLICATION NUMBER: 60/088722
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088738
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088824
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088825
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088863
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876

PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089090
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089908

Query Match
 Best Local Similarity 44.6%; Score 50; DB 12; Length 525;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 LSTEMSPCSVT 21
 DB 441 LAQEMSPCTVT 451

RESULT 10
 US-09-800-729-89
 Sequence 89, Application US/09800729
 Patent No. US2002068319A1
 GENERAL INFORMATION:
 APPLICANT: NI et al.
 TITLE OF INVENTION: 32 Human secreted proteins
 FILE REFERENCE: PZ044P1
 CURRENT APPLICATION NUMBER: US/09/800,729
 PRIOR FILING DATE: 2001-03-08
 PRIOR APPLICATION NUMBER: PCT/US00/26013
 PRIOR FILING DATE: 2000-09-22
 PRIOR APPLICATION NUMBER: 60/155,709
 PRIOR FILING DATE: 1999-09-24
 NUMBER OF SEQ ID NOS: 217
 SOFTWARE: Patentln Ver. 2.0
 SEQ ID NO 89
 LENGTH: 1745
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-800-729-89

Query Match
 Best Local Similarity 44.6%; Score 50; DB 10; Length 1745;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 LSTEMSPCSVT 21
 DB 424 LAQEMSPCTVT 434

RESULT 11
 US-10-044-807-2
 Sequence 2, Application US/10044807
 Patent No. US20020165187A1
 GENERAL INFORMATION:
 APPLICANT: Turner, C. Alexander Jr.
 TITLE OF INVENTION: No. US20020165187A1 Human Protease and Polynucleotides Encod
 FILE REFERENCE: LEX-0298-USA
 CURRENT APPLICATION NUMBER: US/10/044,807
 PRIOR FILING DATE: 2002-01-11
 PRIOR APPLICATION NUMBER: US 60/261,684
 PRIOR FILING DATE: 2001-01-12
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2

LENGTH: 1762
TYPE: PRT
ORGANISM: homo sapiens
US-10-044-807-2

Query Match
Best Local Similarity 44.6%; Score 50; DB 9; Length 1762;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 LSTEMSPCSVT 21
Db 441 LAQEMSPCTVT 451

RESULT 12
US-09-764-853-554
Sequence 554, Application US/09764853
Patent No. US20020090672A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P0206
CURRENT APPLICATION NUMBER: US/09/764,853
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 554
LENGTH: 1766
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (533)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-554

Query Match
Best Local Similarity 44.6%; Score 50; DB 10; Length 1766;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 LSTEMSPCSVT 21
Db 445 LAQEMSPCTVT 455

RESULT 13
US-10-036-869-30
Sequence 30, Application US/10036869
Patent No. US20020151516A1
GENERAL INFORMATION:
APPLICANT: Mixson, James A
TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
THERAPY
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Connolly, Bove, Lodge, & Rutz
STREET: 1220 Market Street, P.O. Box 2207
CITY: Wilmington
STATE: Delaware
COUNTRY: U.S.A.
ZIP: 19899
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/036,869
FILING DATE: 29-NO. US20020151516A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/985,526
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/608,845
FILING DATE: 16-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: McMorrow Jr., Robert G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 658-9141
TELEFAX: (302) 658-5613
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-10-036-869-30

Query Match
Best Local Similarity 41.1%; Score 46; DB 12; Length 63;
Matches 9; Conservative 5; Mismatches 7; Indels 4; Gaps 1;

QY 1 EYLNK---TQNSLSTEMSPCSVT 21
Db 13 VOYRNNEWTVDGSKSPWSSCSVT 37

RESULT 14
US-09-862-027-30
Sequence 30, Application US/09862027
Patent No. US20020142428A1
GENERAL INFORMATION:
APPLICANT: Hodge, Martin R.
TITLE OF INVENTION: NO. US20020142428A1 Kinases and Uses Thereof
FILE REFERENCE: 35800/234862
CURRENT APPLICATION NUMBER: US/09/862,027
CURRENT FILING DATE: 2001-05-21
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 82
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 656
TYPE: PRT
ORGANISM: Mus musculus
US-09-862-027-30

Query Match
Best Local Similarity 41.1%; Score 46; DB 10; Length 656;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 EYLNKQNSLSTEMSPCS 19
Db 570 EHLNPIRENLGRQWKCA 587

RESULT 15
US-10-042-202-11
Sequence 11, Application US/10042202
Patent No. US20020136733A1
GENERAL INFORMATION:
APPLICANT: Adrian Vivian Sinton Hill, Michael AIDOO,
Catherine Elizabeth Margaret Allsopp, Ajit Lalvani, Magdalena
Plebanski, Hilton Carter Whittle,
TITLE OF INVENTION: MALARIA PEPTIDES
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: WENDEROTH, LIND & PONACK, L.L.P.
STREET: 2033 K Street, N.W., Suite 800,
CITY: Washington
STATE: D.C.,
COUNTRY: U.S.A.
ZIP: 20006-1021
COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordpad for Windows 95
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,202
; FILING DATE: 11-Jan-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/714,175
; FILING DATE: 28-JAN-1997
; APPLICATION NUMBER: WO PCT/GB95/26982
; FILING DATE: 30-MAR-1995
; APPLICATION NUMBER: GB 9406492.0
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER: 2002_0026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)-721-8200
; TELEFAX: (202)-721-8250
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-042-202-11

Query Match 40.2%; Score 45; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 YLNKTONSL 11
   |||||
Db 1 YLNKTONSL 9
```

Search completed: December 6, 2002, 12:54:04
Job time : 7.34098 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:41:41 : Search time 12.3934 Seconds
(without alignments)
162.894 Million cell updates/sec

Title: US-09-931-325c-79

Sequence: 1 EYLKINKNSLSTWSPSCSVT 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	96.4	405	2 S05428	circumsporozoite p
2	99	88.4	424	2 A54533	circumsporozoite p
3	99	88.4	442	2 A54529	circumsporozoite p
4	93	83.0	412	1 OZ2QAF	circumsporozoite p
5	91	81.2	388	2 A39756	circumsporozoite p
6	89	79.5	419	1 OZ2QAM	circumsporozoite p
7	82	73.2	351	1 OZ2OKU	circumsporozoite p
8	82	73.2	363	1 OZ2OKM	circumsporozoite p
9	82	73.2	378	1 OZ2OKB	circumsporozoite p
10	82	73.2	401	1 OZ2QAC	circumsporozoite p
11	81	72.3	387	2 D41156	circumsporozoite p
12	81	72.3	387	2 C41156	circumsporozoite p
13	80	71.4	378	1 OZ2OAL	circumsporozoite p
14	80	71.4	398	1 OZ2OAS	circumsporozoite p
15	77	68.8	343	2 A29319	circumsporozoite p
16	77	68.8	367	2 A32068	circumsporozoite p
17	77	68.8	378	1 OZ2QAV	circumsporozoite p
18	77	68.8	386	2 A48571	circumsporozoite p
19	77	68.8	395	2 A41156	circumsporozoite p
20	77	68.8	429	2 A54504	circumsporozoite p
21	77	68.8	485	2 A60610	circumsporozoite p
22	64	57.1	264	2 A44369	circumsporozoite p
23	62	55.4	332	1 OZ2QMB	circumsporozoite p
24	62	55.4	348	1 OZ2QBK	circumsporozoite p
25	62	55.4	367	1 OZ2OMY	circumsporozoite p
26	52	46.4	388	2 JC6164	circumsporozoite p
27	50	44.6	654	2 T32623	hypothetical prote
28	50	44.6	654	2 A96637	hypothetical prote
29	49	43.8	559	2 S04531	thrombospondin-rel

30	49	43.8	574	2 A46283	sporozoite surface
31	48	42.9	209	2 A41342	circumsporozoite p
32	48	42.9	227	2 B41342	circumsporozoite p
33	48	42.9	656	2 T23338	hypothetical prote
34	48	42.9	1444	2 T18856	angiogenesis inhib
35	47.5	42.4	1558	2 C89114	protein C37C3.6a f
36	47.5	42.4	2167	2 T34395	hypothetical prote
37	46	41.1	156	2 C81868	hypothetical prote
38	46	41.1	401	2 S09626	prtb protein - Esc
39	46	41.1	469	1 S29126	properdin precursor
40	46	41.1	551	2 F84567	probable preprotei
41	46	41.1	551	2 A57189	secy protein homol
42	46	41.1	656	2 T49299	receptor interacti
43	46	41.1	1248	2 A47445	reverse gyrase - S
44	45.5	40.6	396	2 T25699	hypothetical prote
45	45	40.2	38	2 A38869	circumsporozoite p

ALIGNMENTS

```

RESULT 1
S05428
C:Species: Plasmodium falciparum
C:Date: 07-Jun-1990 #sequence-revision 07-Jun-1990 #text-change 09-Jun-2000
C:Accession: S05428; A45527; T60657
R:Campbell, J.R.
Nucleic Acids Res. 17, 5854, 1989
A:Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate
A:Reference number: S05428; MUID:89345189; PMID:266895
A:Accession: S05428
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-405 <CAM>
A:Cross-references: EMBL:X15363
R:Caspers, P.; Gentz, R.; Matile, H.; Pink, J.R.; Sinigaglia, F.
Mol. Biochem. Parasitol. 35, 185-190, 1989
A:Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate
A:Reference number: A45527; MUID:89364996; PMID:2671723
A:Accession: A45527
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <CAS>
A:Cross-references: GB:M22982; GB:J04650; NID:q160168; PIDN:AAA29527.1; PID:q160169
R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.
Mol. Biochem. Parasitol. 37, 275-280, 1989
A:Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell
A:Reference number: A60657; MUID:90114334; PMID:2481827
A:Accession: I60657
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 319-336,354-373 <LOC>
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:329-383/Domain: thrombospondin type 1 repeat homology <YHR1>

Query Match 96.4%; Score 108; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 8.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EYLKINKNSLSTWSPSCSVT 21
Db 326 EYLKINKNSLSTWSPSCSVT 345

RESULT 2
A54533
C:Species: Plasmodium falciparum
C:Date: 28-Oct-1994 #sequence-revision 28-Oct-1994 #text-change 09-Jun-2000
C:Accession: A54533
R:del Portillo, H.A.; Nussenzweig, R.S.; Enea, V.
Mol. Biochem. Parasitol. 24, 289-294, 1987

```

```

A>Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.
A:Reference number: A54533; MUID:87315205; PMID:3306573
A:Accession: A54533
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <DEL>
A:Cross-references: GB:A19752; NID:g160216; PIDN:AAA29555.1; PID:g160217
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:348-402/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match      88.4%; Score 99; DB 2; Length 424;
Best Local Similarity 90.0%; Pred. No. 2.1e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EYLKTONSLSTEWSPCSVT 21
   :||| |||||||||
Db 345 QYLNKTONSLSTEWSPCSVT 364

RESULT 3
A:Accession: A54529
A:Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain Wellcome)
C:Species: Plasmodium falciparum
C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: A54529
R:Lockyer, M.J.; Schwarz, R.T.
Mol. Biochem. Parasitol. 22, 101-108, 1987
A>Title: Strain variation in the circumsporozoite protein gene of Plasmodium falciparum
A:Reference number: A54529; MUID:87115616; PMID:3543671
A:Accession: A54529
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-442 <DOC>
A:Cross-references: GB:M15505; NID:g160214; PIDN:AAA29554.1; PID:g160215
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: tandem repeat
F:366-420/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match      88.4%; Score 99; DB 2; Length 442;
Best Local Similarity 90.0%; Pred. No. 2.2e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EYLKTONSLSTEWSPCSVT 21
   :||| |||||||||
Db 363 QYLNKTONSLSTEWSPCSVT 382

RESULT 4
A:Accession: A54520AF
A:Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate IMTW22)
C:Species: Plasmodium falciparum
C>Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jun-2000
C:Accession: A03388
R:Dame, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wiltz, R.A.; Hochmeyer, W.T.
Science 225, 593-599, 1984
A>Title: Structure of the gene encoding the immunodominant surface antigen on the sporozo-
A:Reference number: A03388; MUID:84250215; PMID:6204383
A:Accession: A03388
A:Molecule type: DNA
A:Residues: 1-412 <DM>
A:Cross-references: GB:K02194; NID:g160160; PIDN:AAA29524.1; PID:g160161
A:Experimental source: Clone 768
C:Comment: Residues 1-16 are the probable signal sequence.
C:Comment: There are 41 copies of a 4-residue repeating unit in the middle domain of the sporoz-
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:336-390/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match      83.0%; Score 93; DB 1; Length 412;
Best Local Similarity 80.0%; Pred. No. 1.7e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EYLKTONSLSTEWSPCSVT 21
   :||| |||:::|||||

```

```
DB      333 QYLKKIKNSISTEMSPCSVT   35Z  
RESULT 5  
A39756  
circumsporozoite protein - Plasmodium relictchenowi  
C.Species: Plasmodium relictchenowi  
C.Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999  
C.Accession: A39756  
J.Lal, A.A.; Goldman, I.F.  
J. Biol. Chem. 266, 6686-6689, 1991  
A.Title: Circumsporozoite protein gene from Plasmodium relictchenowi, a chimpanzee malan  
A.Reference number: A39756; MUID:91201303; PMID:2016283  
A.Accession: A39756  
A>Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-388 <LAL>  
A.Cross-references: GB:M60972; NID:g160228; PIDN:AAM29561.1; PID:g160229  
C.Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
F:312-366/Domain: Thrombospondin type 1 repeat homology <THRA>
```

```
Query Match          81.2%; Score 91; DB 2; Length 388;  
Best Local Similarity 80.0%; Pred. No. 3,3e+06;  
Matches    16; Conservative    3; Mismatches     1; Indels       0; Gaps        0;  
  
QY      2 EYLKNIQNLSIETMSPCSVT   21  
II::IIIIIII  
Db      309 EFLLKIQQNLSTTEMSPCSVV   328
```

```
RESULT 6  
OZZQM  
circumsporozoite protein precursor - Plasmodium cynomolg1 (strain Mulligan/NH)  
N.Alternate names: major sporozoitc surface antigen  
C.Species: Plasmodium cynomogi  
C.Date: 30-Sep-1987 #sequence_revizion 30-Sep-1987 #text_change 07-Nov-1997  
C.Accession: B62255  
R.Galinski, M.R.; Arnot, D.E.; Cochrane, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enne  
Cell 48, 311-319, 1987  
A>Title: The circumsporzoite gene of the Plasmodium cynomolg1 complex.  
A.Reference number: A08889; MUID:87102878; PMID:3802196  
A.Accession: B26255  
A.Molecule type: DNA  
A.Residues: 1-419 <GAL>  
C.Comment: There are three distinct regions in the mature circumsporozoite protein, t  
obic membrane anchoring sequence.  
C.Comment: There are 53 tandem copies of a 4-residue repeat.  
C.Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
C.Keywords: sporozoite; surface antigen; tandem repeat  
F:1-19/DomaIn: signal sequence #status predicted <SIG>  
F:20-419/Product: circumsporozoite protein #status predicted <MAT>  
F:99-310/Region: 4-residue repeats  
F:344-397/DomaIn: thrombospondin type 1 repeat homology <THRA>
```

```
Query Match          79.5%; Score 89; DB 1; Length 419;  
Best Local Similarity 75.0%; Pred.No. 7.3e+06;  
Matches    15; Conservative     5; Mismatches     0; Indels       0; Gaps        0;  
  
QY      2 EYLKNISLIETMSPCSVT   21  
II::IIIIIII  
Db      341 EYLDKIRITSTEMSPCSVV   360
```

```
RESULT 7  
OZZOKU  
circumsporozoite protein precursor - Plasmodium knowlesi (strain Nur1)  
N.Alternate names: sporozoite surface protein  
C.Species: Plasmodium knowlesi  
C.Date: 30-Sep-1987 #sequence_revison 30-Sep-1987 #text_change 16-Jul-1999  
C.Accession: A26253  
R.Sharma, S.; Sveic, P.; Mitchell, G.H.; Godson, G.N.  
Science 229, 779-782, 1985  
A>Title: Diversity of circumsporozoite antigen genes from two strains of the malarial
```

A:Reference number: A26253; MUID:85272582; PMID:4023712
A:Accession: A26253
A:Molecule type: DNA
A:Residues: 1351 <SHA>
C:Cross-references: GB:M1031; NID:g160197; PID:AAA29540.1; PID:g160198
C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-351/Product: circumsporozoite protein #status predicted <MAT>
F:98-223/Region: 9-residue repeats
F:224-241/Region: 9-residue repeats
F:276-329/Domain: thrombospondin type 1 repeat homology <THRI>
Query Match 73.2%; Score 82; DB 1; Length 351;
Best Local Similarity 65.0%; Pred. NO. 7.2e-05;
Matches 13; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
Db 273 DYLKIRNSTEWSPCSVT 292

RESULT 8
OZZQAC
circumsporozoite protein - Plasmodium knowlesi (strain H)
C:Species: Plasmodium knowlesi
C:Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 16-Jul-1999
C:Accession: A90841; A93315; A03389
R:Oakli, L.S.; Svec, P.; Nussenzweig, R.S.; Nussenzweig, V.; Godson, G.N.
Cell 34, 815-822, 1983
A:Title: Structure of the Plasmodium knowlesi gene coding for the circumsporozoite protein
A:Reference number: A90841; MUID:84026486; PMID:6313209
A:Molecule type: DNA
A:Residues: 1-363 <OZA>
A:Cross-references: GB:R00822; NID:g160195; PID:AAA19699.1; PID:g160196
R:Godson, G.N.; Ellis, J.; Svec, P.; Schlesinger, D.H.; Nussenzweig, V.
Nature 305, 29-33, 1983
A:Title: Identification and chemical synthesis of a tandemly repeated immunogenic region
A:Reference number: A93315; MUID:83297689; PMID:6193427
A:Accession: A93315
A:Molecule type: mRNA
A:Residues: 107-208 <GOD>
A:Cross-references: GB:R00772
C:Comment: At least 12 copies of a 12-residue repeating unit occur in this surface protein. Mosquito to the vertebrate host.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:288-341/Domain: thrombospondin type 1 repeat homology <THRI>
Query Match 73.2%; Score 82; DB 1; Length 363;
Best Local Similarity 65.0%; Pred. NO. 7.4e-05;
Matches 13; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
Db 285 DYLKIRNSTEWSPCSVT 304

RESULT 9
OZZQAB
circumsporozoite protein precursor - Plasmodium cynomolgi (strain Berok)
N:Alternate names: major sporozoite surface antigen
C:Species: Plasmodium cynomolgi
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
C:Accession: D26255
R:Galinski, M.R.; Arnott, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea, Cell 48, 311-319, 1987
A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
A:Reference number: A90889; MUID:87102878; PMID:3802196
A:Accession: D26255

A:Molecule type: DNA
A:Residues: 1-378 <GAL>
C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-378/Product: circumsporozoite protein #status predicted <MAT>
F:97-192/Region: 9-residue repeats
F:193-268/Region: 16-residue repeats
F:303-356/Domain: thrombospondin type 1 repeat homology <THRI>
Query Match 73.2%; Score 82; DB 1; Length 378;
Best Local Similarity 70.0%; Pred. NO. 7.8e-05;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Db 300 EYLDKIRSTLGEWSPCSVT 319

RESULT 10
OZZQAC
circumsporozoite protein precursor - Plasmodium cynomolgi (strain Gombak)
N:Alternate names: major sporozoite surface antigen
C:Species: Plasmodium cynomolgi
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
C:Accession: E26255
R:Galinski, M.R.; Arnott, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea Cell 48, 311-319, 1987
A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
A:Reference number: A90889; MUID:87102878; PMID:3802196
A:Accession: E26255
A:Molecule type: DNA
A:Residues: 1-401 <GAL>
A:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-401/Product: circumsporozoite protein #status predicted <MAT>
F:98-278/Region: 11-residue repeats
F:326-379/Domain: thrombospondin type 1 repeat homology <THRI>
Query Match 73.2%; Score 82; DB 1; Length 401;
Best Local Similarity 70.0%; Pred. NO. 8.3e-05;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Db 323 EYLDKIRSTLGEWSPCSVT 342

RESULT 11
D41156
circumsporozoite protein - Plasmodium vivax (isolates B19-2 and P4/B)
C:Species: Plasmodium vivax
C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 01-Dec-1995
C:Accession: D41156; B41156
R:Gari, S.H.; Goldman, I.F.; Poyou, M.M.; Oliveira, S.; Alpers, M.P.; Lal, A.A. J. Biol. Chem. 266, 16297-16300, 1991
A:Title: Wide distribution of the variant form of the human malaria parasite Plasmodium falciparum
A:Reference number: A41156; MUID:91358402; PMID:1885563
A:Accession: D41156
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-387 <QAR>
A:Cross-references: GB:M69061
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:312-365/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 72.3%; Score 81; DB 2; Length 387;
Best Local Similarity 61.9%; Pred. No. 0.00011;
Matches 13; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 EYLKIONSLSTEWSPCSVT 21
|||||: : |||: |||
Db 308 EYLDKVRATVGTWTPCSVT 328

RESULT 12

C41156
Circumsporozoite protein - Plasmodium vivax (isolate B7-4)
C:Species: Plasmodium vivax
C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 01-Dec-1995
C:Accession: C41156
R:Gail, S.H.; Goldman, I.F.; Povoa, M.M.; Oliveira, S.; Alpers, M.P.; Lal, A.A.
J. Biol. Chem. 266, 16297-16300, 1991
A:Title: Wide distribution of the variant form of the human malaria parasite Plasmodium
A:Reference number: A41156; MUID:91358402; PMID:1885563
F:312-365/Domain: thrombospondin type 1 repeat homology <THR1>
Accession: C41156
Status: preliminary
Molecule type: DNA

A:Residues: 1-387 <QAR>
A:Cross-references: GB:M69062
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:312-365/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 72.3%; Score 81; DB 2; Length 387;
Best Local Similarity 61.9%; Pred. No. 0.00011;
Matches 13; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 EYLKIONSLSTEWSPCSVT 21
|||||: : |||: |||
Db 308 EYLDKVRATVGTWTPCSVT 328

RESULT 13

O2ZQIL
Circumsporozoite protein precursor - Plasmodium cynomolgi (strain London)
N:Alternate names: major sporozoite surface antigen
C:Species: Plasmodium cynomolgi
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
C:Accession: A26255
R:Galinski, M.R.; Arnott, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea, Cell 48, 311-319, 1987
A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
A:Reference number: A90889; MUID:87102878; PMID:3802196
Accession: A26255
Molecule type: DNA

A:Residues: 1-378 <GAL>
C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-378/Product: circumsporozoite protein #status predicted <MAT>
F:98-211/Region: 6-residue repeats
F:212-277/Region: 11-residue repeats
F:303-356/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 71.4%; Score 80; DB 1; Length 378;
Best Local Similarity 65.0%; Pred. No. 0.00016;
Matches 13; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 2 EYLKIONSLSTEWSPCSVT 21
|||||: : |||: |||
Db 300 EYLDKIRSTIGVWSPCSVT 319

RESULT 14

O2ZQNS
Circumsporozoite protein precursor - Plasmodium cynomolgi (strain Ceylon)

N:Alternate names: major sporozoite surface antigen
C:Species: Plasmodium cynomolgi
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
C:Accession: C26255

R:Galinski, M.R.; Arnott, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea Cell 48, 311-319, 1987
A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
A:Reference number: A90889; MUID:87102878; PMID:3802196
Accession: C26255
Molecule type: DNA
A:Residues: 1-398 <GAL>
C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-398/Product: circumsporozoite protein #status predicted <MAT>
F:97-240/Region: 9-residue repeats
F:241-291/Region: 17-residue repeats
F:323-376/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 71.4%; Score 80; DB 1; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.00017;
Matches 13; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 2 EYLKIONSLSTEWSPCSVT 21
|||||: : |||: |||
Db 320 EYLDKIRSTIGVWSPCSVT 339

RESULT 15

A29319
Circumsporozoite protein - Plasmodium vivax (strain Sal-I) (fragment)
C:Species: Plasmodium vivax
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 17-Mar-1999
C:Accession: A29319; A44187
R:de la Cruz, V.F.; Lal, A.A.; Welsh, J.A.; McCutchan, T.F.
J. Biol. Chem. 262, 6464-6467, 1987
A:Title: Evolution of the immunodominant domain of the circumsporozoite protein gene
A:Reference number: A29319; MUID:87194878; PMID:2437120
Accession: A29319
Molecule type: DNA

A:Residues: 1-343
R:McCutchan, T.F.; Lal, A.A.; de la Cruz, V.F.; Miller, L.H.; Maloy, W.L.; Charoenvit Science 230, 1381-1383, 1985
A:Title: Sequence of the immunodominant epitope for the surface protein on sporozoite
A:Reference number: A44187; MUID:86070222; PMID:2416057
Accession: A44187
Molecule type: DNA
A:Status: preliminary
A:Residues: 1-98 <KMC>
A:Cross-references: GB:J02751
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:268-321/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 68.8%; Score 77; DB 2; Length 343;
Best Local Similarity 60.0%; Pred. No. 0.00041;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 2 EYLKIONSLSTEWSPCSVT 21
|||||: : |||: |||
Db 265 EYLDKVRATVGTWTPCSVT 284

Search completed: December 6, 2002, 12:46:03
Job time : 13.3934 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:38:55 : Search time 6.88525 Seconds
(without alignments)
126.503 Million cell updates/sec

Title: US-09-931-325c-79
Perfect score: 112
Sequence: 1 IEXLINKIONSLESTFSPSCSVT 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	96.4	397	1	CSP_PLAFO
2	99	88.4	424	1	CSP_PLAFT
3	99	88.4	442	1	CSP_PLAFW
4	93	83.0	412	1	CSP_PLAFR
5	89	81.2	388	1	CSP_PLARE
6	89	79.5	419	1	CSP_PLACM
7	82	73.2	351	1	CSP_PLAKU
8	82	73.2	363	1	CSP_PLAKH
9	82	73.2	378	1	CSP_PLAKB
10	80	71.4	398	1	CSP_PLACC
11	79	70.5	401	1	CSP_PLACG
12	77	68.8	343	1	CSP_PLAVS
13	77	68.8	378	1	CSP_PLACL
14	77	68.8	378	1	CSP_PLAVB
15	77	68.8	366	1	CSP_PLASI
16	77	68.8	393	1	CSP_PLARR
17	77	68.8	429	1	CSP_PLAMA
18	62	55.4	339	1	CSP_PLABE
19	62	55.4	347	1	CSP_PLABA
20	62	55.4	367	1	CSP_PLAYO
21	49	43.8	559	1	TRAP_PLAFA
22	48	42.9	1435	1	TR85_HUMAN
23	46	41.1	401	1	TISP_ECOLI
24	46	41.1	469	1	PROP_HUMAN
25	46	41.1	551	1	SECY_ARATH
26	46	41.1	656	1	RIKI_MOUSE
27	46	41.1	1248	1	TOPG_STULAC
28	45	40.2	191	1	EXBB_MAIZE
29	45	40.2	247	1	CDH_ENTRL
30	45	40.2	269	1	EXBA_MAIZE
31	45	40.2	307	1	RA31_SCHPO
32	45	40.2	553	1	SECY_MAIZE
33	45	40.2	1074	1	SM5A_HUMAN

34	45	40.2	1077	1	SM5A_MOUSE
35	44	39.3	450	1	TH12_YEAST
36	44	39.3	639	1	CA1C_RABIT
37	44	39.3	800	1	PT27_YEAST
38	44	39.3	849	1	VNCS_AEDEV
39	44	39.3	1173	1	TSPI_XENIA
40	44	39.3	3063	1	CA1C_HUMAN
41	43.5	38.8	2052	1	MY10_BOVIN
42	43	38.4	141	1	V192_FOMPV
43	43	38.4	209	1	RNF6_VIRCH
44	43	38.4	229	1	NHBI_RHOB
45	43	38.4	272	1	VAL1_MSVK

ALIGNMENTS

RESULT 1
CSP_PLAFO STANDARD; PRT; 397 AA.
AC P19597; Q25798; 17, Created)
DT 01-FEB-1991 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69345189; PubMed=2668895;
RA Campbell J.R.;
RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria
candidate vaccine antigen."
RL Nucleic Acids Res. 17:5854-5854(1989).
RN [2]
RP REVISIONS.
RA Campbell J.R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92155298; PubMed=1346766;
RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,
Thomas A.W., Bagar S., Cochran M.A., Thanassi J., Levine M.M.,
Hackett C.S.;
RT "Plasmodium falciparum: in vitro characterization and human
infectivity of a cloned line."
RL Exp. Parasitol. 74:159-168(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89364986; PubMed=2671723;
RA Caspers P., Gentz R., Matile H., Pink J.R., Sinigaglia F.;
RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum
isolate used in malaria vaccine trials."
ML Biochem. Parasitol. 35:185-190(1989).
-1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
VERTERATE HOST).
-1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
-1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: X15363; CAA33421.1; -

```

DR EMBL; M83886; AAA29521.1; -.
DR EMBL; M22982; AAA29527.1; -.
DR PIR; S05428; S05428.
DR PIR; A45527; A45527.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 1.
DR SMART; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 101 272 43 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT CONFILCT 194 194 A -> ANPNANPNA (IN REF. 4).
SQ SEQUENCE 397 AA; 42646 MW; 9E81146F59EBCEA3 CRC64;

Query Match 96.4%; Score 108; DB 1; Length 397;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

2 EYLKIONSLESTWSPSCSVT 21

```

DB 318 EYLKIONSLESTWSPSCSVT 337
|||||

```

```

RESULT 2
CSP_PLAFT STANDARD: PRT: 424 AA.
ID CSP_PLAFT 13814:
AC P13814:
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate t4 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5846;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S., Enea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from Thailand";
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
-----
DR EMBL; M19752; AAA29555.1; -.
DR PIR; A54533; A54533.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 424 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;

```

```

Query Match 88.4%; Score 99; DB 1; Length 424;
Best Local Similarity 90.0%; Pred. No. 1.4e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

2 EYLKIONSLESTWSPSCSVT 21

```

DB 345 EYLKIONSLESTWSPSCSVT 364
|||||

```

RESULT 3

```

CSP_PLAFT STANDARD: PRT: 442 AA.
ID CSP_PLAFT P08307:
AC P08307:
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87115616; PubMed=3543671;
RA Lockyer M.J., Schwarz R.T.;
RT "Strain variation in the circumsporozoite protein gene of Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 22:101-108(1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
-----
DR EMBL; M15505; AAA29554.1; -.
DR PIR; A54529; A54529.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 CIRCUMSPOROZOITE PROTEIN.
FT CHAIN 17 442 47 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT DOMAIN 130 320 BD57A9A152B85E03 CRC64;
SQ SEQUENCE 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;

Query Match 88.4%; Score 99; DB 1; Length 442;
Best Local Similarity 90.0%; Pred. No. 1.5e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

2 EYLKIONSLESTWSPSCSVT 21

```

DB 363 EYLKIONSLESTWSPSCSVT 382
|||||

```

RESULT 4

```

CSP_PLAFT STANDARD: PRT: 412 AA.
ID CSP_PLAFT P02893:
AC P02893:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE  Circumsporozoite protein precursor (CS) .
OS  Plasmodium falciparum.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5833;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=84250215; PubMed=6204383;
RA  Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA  Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA  Sanders G.S., Keddy E.P., Diggs C.L., Miller L.H.;
RT  "Structure of the gene encoding the immunodominant surface antigen on
RT  the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL  Science 225:593-599(1984).
CC  -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC  SURFACE ANTIGEN ON THE SPOOROITE (THE INFECTIVE STAGE OF THE
CC  MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC  VERTEBRATE HOST).
CC  -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC  ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC  WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC  -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; K02194; AAA29524.1; -
DR  PIR; A03388; OZQAF.
DR  InterPro; IPR003067; Circmsprzoite.
DR  InterPro; IPR000884; TSP1.
DR  Pfam; PF00090; tsp_1; 1.
DR  PRINTS; PR01303; CRCMSPRZOITE.
DR  SMART; SM00209; TSP1; 1.
KW  Malaria; Sporozoite; Repeat; Signal.
FT  SIGNAL 1 16 PROBABLE.
FT  CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.
FT  DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.
SQ  SEQUENCE 412 AA; 44420 MW; 1EEED3DE9095F8 CRC64;

Query Match 83.0%; Score 93; DB 1; Length 412;
Best Local Similarity 80.0%; Pred. No. 1,1e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

CY 2 EYLKIONSLSTWSPSCSVT 21
Db 333 QYLKKIKNSISTWSPSCSVT 352

RESULT 5
CSP_PLARE STANDARD; PRT; 388 AA.
AC P26694;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91201303; PubMed=2016283;
RA Lal A.A., Goldman I.F.;
RT "Circumsporozoite protein gene from Plasmodium reichenowi, a
RT chimpanzee malaria parasite evolutionarily related to the human
RT malaria parasite Plasmodium falciparum.";
RL J. Biol. Chem. 266:6686-6689(1991).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROITE (THE INFECTIVE STAGE OF THE

```

```

CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; M60972; AAA29561.1; -
DR  PIR; A39756; A39756.
DR  InterPro; IPR003067; Circmsprzoite.
DR  InterPro; IPR000884; TSP1.
DR  Pfam; PF00090; tsp_1; 1.
DR  PRINTS; PR01303; CRCMSPRZOITE.
DR  SMART; SM00209; TSP1; 1.
KW  Malaria; Sporozoite; Repeat; Signal.
FT  SIGNAL 1 16 PROBABLE.
FT  CHAIN 17 388 CIRCUMSPOROZOITE PROTEIN.
FT  DOMAIN 120 267 62 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ  SEQUENCE 388 AA; 42245 MW; C031EEFBE2E35604 CRC64;

Query Match 81.2%; Score 91; DB 1; Length 388;
Best Local Similarity 80.0%; Pred. No. 2,1e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

CY 2 EYLKIONSLSTWSPSCSVT 21
Db 309 EFLKQIONNLSSTWSPSCSVT 328

RESULT 6
CSP_PLACM STANDARD; PRT; 419 AA.
AC P08676;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium cynomolgi (Strain Mulligan/NIH).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5832;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochran A.H., Barnwell J.W.,
RA Nussenzweig R.S., Enea V.;
RT "The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; M15102; AAA29539.1; -

```

```
DR PIR: B26255; OZ20AM.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; TSP_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 419 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 99 314 54 X 4 AA TANDEM REPEATS OF N-A-[DG]-G.
SQ SEQUENCE 419 AA; 38924 MM; 8F46CDD8A1BAEFF4 CRC64;

Query Match 79.5%; Score 89; DB 1; Length 419;
Best Local Similarity 75.0%; Pred. NO. 4.5e-06;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 2 EYLKIONSLSIEWSPCSVT 21
DB 341 EYLKIRSRSTIEWSPCSVT 360

SULT 7
CSP_PLAKU STANDARD; PRT; 351 AA.
AC P04922;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium knowlesi (strain nurl).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5852;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85272582; PubMed=4023712;
RA Sharma S., Svec P., Mitchell G.H., Godson G.N.;
RT "Diversity of circumsporozoite antigen genes from two strains of the
RL malarial parasite Plasmodium knowlesi.";
Science 229:779-782(1985).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTERATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: M1031; AAA29540.1; -.
DR EMBL: A26253; OZ20KU.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; TSP_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 351 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 101 235 15 X 9 AA TANDEM REPEATS OF A-A-G-A-G-G-
SQ SEQUENCE 351 AA; 34782 MM; A85E87A152E6485B CRC64;

Query Match 73.2%; Score 82; DB 1; Length 351;
Best Local Similarity 65.0%; Pred. NO. 4.2e-05;
Matches 13; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 2 EYLKIONSLSIEWSPCSVT 21
DB 273 EYLKIRSRSTIEWTPCSVT 292

RESULT 8
CSP_PLAKH STANDARD; PRT; 363 AA.
ID CSP_PLAKH
AC P02894;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium knowlesi (strain H).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5851;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84026486; PubMed=6313209;
RA Ozaki L.S., Svec P., Nussenzweig R.S., Nussenzweig V., Godson G.N.;
RT "Structure of the plasmodium knowlesi gene coding for the
RL circumsporozoite protein.";
Cell 34:815-822(1983).
RN [2]
RP SEQUENCE OF 84-258 FROM N.A.
RX MEDLINE=83297689; PubMed=6193427;
RA Godson G.N., Ellis J., Svec P., Schlesinger D.H., Nussenzweig V.;
RT "Identification and chemical synthesis of a tandemly repeated
RL immunogenic region of Plasmodium knowlesi circumsporozoite protein.";
Nature 305:29-33(1983).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTERATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: K00822; AAA19699.1; -.
DR EMBL: K00772; AAA29556.1; -.
DR PIR: A03389; OZ20AK.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; TSP_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 363 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 241 12 X 12 AA TANDEM REPEATS OF N-A-G-Q-P-Q-
SQ SEQUENCE 363 AA; 36793 MM; 574DF4BD320A7955 CRC64;

Query Match 73.2%; Score 82; DB 1; Length 363;
Best Local Similarity 65.0%; Pred. NO. 4.4e-05;
Matches 13; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
```

```

RESULT 9
CSP_PLACC STANDARD; PRT: 378 AA.
AC P08672;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium cynomolgi (strain Berek).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
RA Nussenzweig R.S., Enea V.;
RT "The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- DOMAIN: THERE ARE 10 TANDEM COPIES OF A 9-RESIDUE REPEAT (PRECEDED BY A 6-RESIDUE INCOMPLETE REPEAT) AND 3 TANDEM COPIES OF A 16-RESIDUE REPEAT (FOLLOWED BY 3 SHORTER, INCOMPLETE COPIES).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M15104; AAA29532.1; -
DR PIR: D62625; OZQOAS.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF000090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 378 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 97 192 10.5 X 9 AA REPEATS.
FT DOMAIN 193 240 3 X 16 AA TANDEM REPEATS.
FT REPEAT 241 251
FT REPEAT 252 260
FT REPEAT 261 268
SQ SEQUENCE 378 AA; 36286 MW; 779BA081C140793F CRC64;

Query Match 73.2%; Score 82; DB 1; Length 378;
Best Local Similarity 70.0%; Pred. NO. 4.6e-05;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 EYLNKIONSLSTEWSPCSVT 21
Db 300 EYLDKIRSTIGVWSPCSVT 319

```

```

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5829;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
RA Nussenzweig R.S., Enea V.;
RT "The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M15103; AAA29533.1; -
DR PIR: C26255; OZQOAS.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF000090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 398 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 97 249 17 X 9 AA TANDEM REPEATS OF A-G-N-N-A-A-A-G-E.
SQ SEQUENCE 398 AA; 37718 MW; 6DFA2E8A62E055F CRC64;

Query Match 71.4%; Score 80; DB 1; Length 398;
Best Local Similarity 65.0%; Pred. NO. 9.6e-05;
Matches 13; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 EYLNKIONSLSTEWSPCSVT 21
Db 320 EYLDKIRSTIGVWSPCSVT 339

RESULT 11
CSP_PLACC STANDARD; PRT: 401 AA.
AC P08674;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium cynomolgi (strain Gombak).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5830;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
RA Nussenzweig R.S., Enea V.;
RT "The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES

```

CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M15100; AAA29536.1; -
DR PIR: E26255; OZ2QAC.
DR InterPro: IPR003067; Circmsprzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
KM Malaria: Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 401 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 278 17 X 11 AA TANDEM REPEATS OF [DG]-G-A-A-
A-A-G-G-G-G-N.
SQ SEQUENCE 401 AA: 36664 MW: 57D66268238503E CRC64:
Query Match 70.5%; Score 79; DB 1; Length 401;
Best Local Similarity 65.0%; Pred. No. 0.00014;
Matches 13; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
OY 2 EYLKIONSLEWSPSCSVT 21
DB 323 EYLKIRSTLGVWSPCTVT 342
RESULT 12
CSP_PLAYS STANDARD: PRT: 343 AA.
AC P13826;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein (CS) (Fragment).
OS Plasmodium vivax (strain Salvador I).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=126793;
[1]
RX SEQUENCE FROM N.A. PubMed=2416057;
MEDLINE=86070222; PubMed=2416057;
McCutchan T.F., Lal A.A., de la Cruz V.F., Miller L.H., Maloy W.L.,
Chaoenwatt T., Beaudoin R.L., Guerry P., Wistlar R. Jr., Hoffman S.L.,
Hockmeyer W.T., Collins W.E., Wirth D.;
"Sequence of the immunodominant epitope for the surface protein on
RT sporozoites of Plasmodium vivax.";
RL Science 230:1381-1383(1985).
[2]
RP SEQUENCE FROM N.A. PubMed=2437120;
MEDLINE=87194878; PubMed=2437120;
de la Cruz V.F., Lal A.A., Welsh J.A., McCutchan T.F.;
"Evolution of the immunodominant domain of the circumsporozoite
RT protein gene from Plasmodium vivax. Implications for vaccines.";
RL J. Biol. Chem. 262:6464-6467(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.

KM Sporozoite; Malaria; Repeat.
FT NON_TER 1 19 X 9 AA TANDEM REPEATS OF P-G-D-R-A-D-
FT DOMAIN 63 243 G-O-P.
SQ SEQUENCE 343 AA: 34155 MW: 308EFP5B8C150FC3 CRC64:
Query Match 68.8%; Score 77; DB 1; Length 343;
Best Local Similarity 60.0%; Pred. No. 0.00023;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
OY 2 EYLKIONSLEWSPSCSVT 21
DB 265 EYLKVRATVGTWTPCSTV 284
RESULT 13
CSP_PLACL STANDARD: PRT: 378 AA.
AC P08675;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium cynomolgi (strain London).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5831;
[1]
RX SEQUENCE FROM N.A. PubMed=3802196;
MEDLINE=87102878; PubMed=3802196;
Galinski M.R., Arnot D.E., Cochran A.H., Barnwell J.W.,
Nussenzweig R.S., Enea V.;
"The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M15101; AAA29537.1; -
DR PIR: A26255; OZ2QAL.
DR InterPro: IPR003067; Circmsprzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
KM Malaria: Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 378 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 211 18 X 6 AA TANDEM REPEATS OF D-G-A-R-A-
[EA].
FT DOMAIN 212 277 6 X 11 AA TANDEM REPEATS OF G-N-Q-A-G-G-
O-A-G-A-G.
SQ SEQUENCE 378 AA: 37462 MW: 8295A913C36420C5 CRC64:
Query Match 68.8%; Score 77; DB 1; Length 378;
Best Local Similarity 60.0%; Pred. No. 0.00026;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
OY 2 EYLKIONSLEWSPSCSVT 21
DB 300 EYLKIRSTLGVWSPCTVT 319

```
RESULT 14
CSP_PLAVB          STANDARD;          PRT;          378 AA.
ID CSP_PLAVB
AC P08677;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium vivax (strain Belen).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86044510; Pubmed=241847;
RA Annot D.E., Barnwell J.W., Tam J.P., Nussenzweig V., Nussenzweig R.S.,
RA Enea V.;
RT "Circumsporozoite protein of Plasmodium vivax: gene cloning and
RT characterization of the immunodominant epitope.";
RL Science 230:815-818(1985).
RN [2]
RP REVISIONS.
RX MEDLINE=89042133; Pubmed=3054880;
RA Annot D.E., Barnwell J.W., Stewart M.J.;
RT "Does biased gene conversion influence polymorphism in the
RT circumsporozoite protein-encoding gene of Plasmodium vivax?";
RL Proc. Natl. Acad. Sci. U.S.A. 85:8102-8106(1988).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M1926; AAA29526.1; -; ALT_SEQ.
DR EMBL; J02751; AAA29529.1; -; ALT_SEQ.
DR PIR; A26256; OZQAV.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF000090; tsp_1; 1.
DR Pfam; PF000084; TSP1.
DR PRINTS; PRO1303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1. 1.
DR PROSITE; PS50092; TSP1. 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 97 267
FT DOMAIN 97 267
FT 19 X 9 AA TANDEM REPEATS OF P-G-D-R-A-D-
FT G-Q-P.
FT CONFLICT 36 36 G->E (IN REF. 1).
FT CONFLICT 96 96 G->R (IN REF. 1).
FT CONFLICT 295 295 E->A (IN REF. 1).
FT CONFLICT 328 328 R->S (IN REF. 1).
SQ SEQUENCE 378 AA; 37800 MW; C84B5BED05E3C9ED CRC64;

Query Match 68.8%; Score 77; DB 1; Length 378;
Best Local Similarity 60.0%; Pred. No. 0.00026;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
```

```
RESULT 15
CSP_PLASI          STANDARD;          PRT;          386 AA.
ID CSP_PLASI
AC 003110;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium simium.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5859;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93149205; Pubmed=8426613;
RA Goldman I.F., Gari S.H., Millet P.G., Collins W.E., Lal A.A.;
RT "Circumsporozoite protein gene of Plasmodium simium, a Plasmodium
RT vivax-like monkey malaria parasite.";
RL Mol. Biochem. Parasitol. 57:177-180(1993).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L05068; AAA29525.1; -;
DR PIR; A48571; AAA29525.1; -;
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF000090; tsp_1; 1.
DR Pfam; PF000084; TSP1.
DR PRINTS; PRO1303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1. 1.
DR PROSITE; PS50092; TSP1. 1.
DR Malaria; Sporozoite; Signal; Repeat.
FT SIGNAL 1 19
FT CHAIN 20 386
FT DOMAIN 96 275
FT 20 X 9 AA TANDEM REPEATS OF G-D-R-A-[AD]-
FT G-Q-P-A.
SQ SEQUENCE 386 AA; 38567 MW; A0097D4BDE5548DB CRC64;

Query Match 68.8%; Score 77; DB 1; Length 386;
Best Local Similarity 60.0%; Pred. No. 0.00026;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
```

Search completed: December 6, 2002, 12:44:02
Job time : 7.88525 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:41:26 : Search time 24.0984 Seconds
(without alignments)
179.555 Million cell updates/sec

Title: US-09-931-325C-79

Perfect score: 112

Sequence: 1 EYLNKIONSLSSTWSPCSVT 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	108	96.4	80	5	Q9U0P6	Q9U0P6 plasmodium
2	106	94.6	106	5	Q9U0P0	Q9U0P0 plasmodium
3	102	91.1	79	5	Q9U0Q2	Q9U0Q2 plasmodium
4	102	91.1	115	5	Q25835	Q25835 plasmodium
5	102	91.1	436	5	Q27325	Q27325 plasmodium
6	100	89.3	420	5	Q25838	Q25838 plasmodium
7	99	88.4	80	5	Q9U0P8	Q9U0P8 plasmodium
8	99	88.4	80	5	Q9U0P3	Q9U0P3 plasmodium
9	99	88.4	80	5	Q9U0P2	Q9U0P2 plasmodium
10	99	88.4	80	5	Q9U0P0	Q9U0P0 plasmodium
11	99	88.4	80	5	Q9U0P9	Q9U0P9 plasmodium
12	99	88.4	80	5	Q9U0P9	Q9U0P9 plasmodium
13	99	88.4	115	5	Q25836	Q25836 plasmodium
14	99	88.4	115	5	Q9U0P7	Q9U0P7 plasmodium
15	99	88.4	117	5	Q25795	Q25795 plasmodium
16	99	88.4	117	5	Q25795	Q25795 plasmodium

17	99	88.4	117	5	Q25796	Q25796 plasmodium
18	99	88.4	117	5	Q25797	Q25797 plasmodium
19	99	88.4	408	5	Q25729	Q25729 plasmodium
20	99	88.4	416	5	Q25829	Q25829 plasmodium
21	99	88.4	424	5	Q99256	Q99256 plasmodium
22	99	88.4	424	5	Q27425	Q27425 plasmodium
23	99	88.4	432	5	Q25827	Q25827 plasmodium
24	99	88.4	432	5	Q27246	Q27246 plasmodium
25	99	88.4	442	5	Q25830	Q25830 plasmodium
26	97	86.6	80	5	Q9U0P9	Q9U0P9 plasmodium
27	95	84.8	80	5	Q9U0Q4	Q9U0Q4 plasmodium
28	95	84.8	80	5	Q9U0P1	Q9U0P1 plasmodium
29	95	84.8	80	5	Q9U0P4	Q9U0P4 plasmodium
30	95	84.8	117	5	Q25794	Q25794 plasmodium
31	94	83.9	106	5	Q9U0P2	Q9U0P2 plasmodium
32	93	83.0	80	5	Q9U0P0	Q9U0P0 plasmodium
33	93	83.0	80	5	Q9U0P7	Q9U0P7 plasmodium
34	93	83.0	80	5	Q9U0P4	Q9U0P4 plasmodium
35	93	83.0	80	5	Q9U0P9	Q9U0P9 plasmodium
36	92	82.1	73	5	Q25833	Q25833 plasmodium
37	91	81.2	80	5	Q9U0P1	Q9U0P1 plasmodium
38	91	81.2	80	5	Q9U0P5	Q9U0P5 plasmodium
39	91	81.2	80	5	Q9U0P1	Q9U0P1 plasmodium
40	91	81.2	115	5	Q25837	Q25837 plasmodium
41	91	81.2	115	5	Q25839	Q25839 plasmodium
42	91	81.2	383	5	Q9GPN1	Q9GPN1 plasmodium
43	91	81.2	420	5	Q25831	Q25831 plasmodium
44	91	81.2	436	5	Q25828	Q25828 plasmodium
45	91	81.2	452	5	Q25834	Q25834 plasmodium

ALIGNMENTS

RESULT 1
ID Q9U0P6 PRELIMINARY; PRT; 80 AA.
AC Q9U0P6.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA264;
RA "de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ269663; CAB64182.1; -
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; csp_L; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1 80
FT NON_TER 80 80
SQ SEQUENCE 80 AA. 9002 MW: 1CERAE08EGC9E976 CRC64;

Query Match 96.4%; Score 108; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 7.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EYLNKIONSLSSTWSPCSVT 21
DB 15 EYLNKIONSLSSTWSPCSVT 34

```

RESULT 2
Q906P0 PRELIMINARY; PRT; 106 AA.
AC Q906P0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NX NCBI_Taxid=5833;
RN [1]
RE SEQUENCE FROM N.A.
RC STRAIN=1068;
RX MEDLINE=20164888; PubMed=10699256;
RA Allouche A., Silveira H., Conway D.J., Bojang K., Doherty T.,
  Cohen J., Pinder M., Greenwood B.M.;
RT "High-throughput sequence typing of P-cell epitope polymorphisms in
  Plasmodium falciparum circumsporozoite protein.";
  Mol. Biochem. Parasitol. 106:273-282(2000).
DR EMBL; AF181835; AAF03136.1; -
  InterPro: IPR003067; Circspprzoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp.1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPL; 1.
FT NON_TER 1
FT NON_TER 106
SQ SEQUENCE 106 AA; 11815 MW; 42E13D4A19BD0D3C CRC64;

Query Match
Best Local Similarity 94.6%; Score 106; DB 5; Length 106;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 EXLNKIONSISTEMSPCSVT 21
DB 43 EXLNKIONSISTEMSPCSVT 62

RESULT 3
Q900Q2 PRELIMINARY; PRT; 79 AA.
AC Q900Q2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NX NCBI_Taxid=5833;
RN [1]
RE SEQUENCE FROM N.A.
RC STRAIN=DS1;
RA de Stricker K., Vuust J., Jepsen S., Oeuviy C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
  falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
  Burmese field isolates and from laboratory strains.";
  Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269945; CAB64243.1; -
  InterPro: IPR003067; Circspprzoite.
DR InterPro: IPR000884; TSPL.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPL; 1.
FT NON_TER 1
FT NON_TER 79
SQ SEQUENCE 79 AA; 8893 MW; 56BFA76D859B416 CRC64;

Query Match
Best Local Similarity 91.1%; Score 102; DB 5; Length 79;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 EXLNKIONSISTEMSPCSVT 21
DB 15 EXLNKIONSISTEMSPCSVT 34

RESULT 4
Q25835 PRELIMINARY; PRT; 115 AA.
AC Q25835;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NX NCBI_Taxid=5833;
RN [1]
RE SEQUENCE FROM N.A.
RC STRAIN=828;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwattanas S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
  falciparum from Thai field isolates.";
  Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83157; AAA29570.1; -
  InterPro: IPR003067; Circspprzoite.
DR InterPro: IPR000884; TSPL.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPL; 1.
FT NON_TER 1
FT NON_TER 115
SQ SEQUENCE 115 AA; 12925 MW; 9BE52CFE12BD48F CRC64;

Query Match
Best Local Similarity 91.1%; Score 102; DB 5; Length 115;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EXLNKIONSISTEMSPCSVT 21
DB 36 EXLNKIONSISTEMSPCSVT 55

RESULT 5
Q27325 PRELIMINARY; PRT; 436 AA.
AC Q27325;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NX NCBI_Taxid=5833;
RN [1]
RE SEQUENCE FROM N.A.
RC STRAIN=837;
RX MEDLINE=84250215; PubMed=6204383;
RA Dane J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
  Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
  Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
  the sporozoite of the human malaria parasite Plasmodium falciparum.";
  Science 225:593-599(1984).
RN [2]
RE SEQUENCE FROM N.A.
RC STRAIN=837;
RA la Cruz V.F.;
RT Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RE SEQUENCE FROM N.A.
RC STRAIN=837;

```

RA Jongwutives S., Tanabe K., Kanbara H.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL: M83164; AAA29542.1; -;
DR EMBL: M83150; AAA29563.1; -;
DR EMBL: M83163; AAA29576.1; -;
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
SQ SEQUENCE 436 AA; 4668 MW; 5B42FF3348B68655 CRC64;

Query Match 91.1%; Score 102; DB 5; Length 436;
Best Local Similarity 95.0%; Pred. No. 4.4e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 EYLNKIONSLSTEWSPCSVT 21
DB 357 EYLNKIONSLSTEWSPCSVT 376

RESULT 6

O25838 PRELIMINARY; PRT; 420 AA.

ID O25838;
AC O25838;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=835B;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RT falciparum from Thai field isolates."
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL: M83161; AAA29574.1; -;
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
SQ SEQUENCE 420 AA; 4515 MW; 3A85B92432C2893C CRC64;

Query Match 89.3%; Score 100; DB 5; Length 420;
Best Local Similarity 90.0%; Pred. No. 8.9e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 EYLNKIONSLSTEWSPCSVT 21
DB 341 EYLNKIONSLSTEWSPCSVT 360

RESULT 7

O900P8 PRELIMINARY; PRT; 80 AA.

ID O900P8;
AC O900P8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Circumsporozoite protein (fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M4;

RA de Stricker K., Vuust J., Jepsen S., Oeuvery C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269957; CAB64237.1; -;
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9073 MW; A8F404B8FB142B1E CRC64;

Query Match 88.4%; Score 99; DB 5; Length 80;
Best Local Similarity 90.0%; Pred. No. 2e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 EYLNKIONSLSTEWSPCSVT 21
DB 15 EYLNKIONSLSTEWSPCSVT 34

RESULT 8

O900P3 PRELIMINARY; PRT; 80 AA.

ID O900P3;
AC O900P3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Circumsporozoite protein (fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D4405;
RA de Stricker K., Vuust J., Jepsen S., Oeuvery C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269959; CAB64188.1; -;
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9033 MW; BA71EBE0DB03193B CRC64;

Query Match 88.4%; Score 99; DB 5; Length 80;
Best Local Similarity 90.0%; Pred. No. 2e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 EYLNKIONSLSTEWSPCSVT 21
DB 15 EYLNKIONSLSTEWSPCSVT 34

RESULT 9

O900P2 PRELIMINARY; PRT; 80 AA.

ID O900P2;
AC O900P2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Circumsporozoite protein (fragment).
GN CS.
OS Plasmodium falciparum.

```
OC Eukaryota: Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D4416;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269970; CAB64189.1; -.
DR InterPro: IPR003067; Circsmpztoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSPL; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9015 MW; 1CF404B8FB142C73 CRC64;

Query Match 88.4%; Score 99; DB 5; Length 80;
Best Local Similarity 90.0%; Pred. No. 2e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 EYLKTIKNSLSTWSPSCSVT 21
DB 15 QYLKTIKNSLSTWSPSCSVT 34

RESULT 10
O9TW00 PRELIMINARY; PRT; 80 AA.
AC O9TW00;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota: Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EUP/SP, AND MAD20;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269953; CAB64178.1; -.
DR EMBL; AJ269950; CAB64175.1; -.
DR InterPro: IPR003067; Circsmpztoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSPL; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9074 MW; 4614F9D18F1C0334 CRC64;

Query Match 88.4%; Score 99; DB 5; Length 80;
Best Local Similarity 90.0%; Pred. No. 2e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 EYLKTIKNSLSTWSPSCSVT 21
DB 15 QYLKTIKNSLSTWSPSCSVT 34

RESULT 11
O9TW00 PRELIMINARY; PRT; 80 AA.
AC O9TW00;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota: Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EUP/SP, AND MAD20;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269953; CAB64178.1; -.
DR EMBL; AJ269950; CAB64175.1; -.
DR InterPro: IPR003067; Circsmpztoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSPL; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9074 MW; 4614F9D18F1C0334 CRC64;

Query Match 88.4%; Score 99; DB 5; Length 80;
Best Local Similarity 90.0%; Pred. No. 2e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 EYLKTIKNSLSTWSPSCSVT 21
DB 15 QYLKTIKNSLSTWSPSCSVT 34

RESULT 12
O9TW00 PRELIMINARY; PRT; 80 AA.
AC O9TW00;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota: Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-M0, M7, AND M1;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269955; CAB64238.1; -.
DR EMBL; AJ269958; CAB64238.1; -.
DR EMBL; AJ269959; CAB64239.1; -.
DR EMBL; AJ269956; CAB64241.1; -.
DR InterPro: IPR003067; Circsmpztoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSPL; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9074 MW; A8F40C90DB1C033E CRC64;

Query Match 88.4%; Score 99; DB 5; Length 80;
Best Local Similarity 90.0%; Pred. No. 2e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota: Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D4393, AND D4372;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269966; CAB64185.1; -.
DR EMBL; AJ269964; CAB64183.1; -.
DR InterPro: IPR003067; Circsmpztoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSPL; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9059 MW; A756D1FPC41C1C21 CRC64;

Query Match 88.4%; Score 99; DB 5; Length 80;
Best Local Similarity 85.0%; Pred. No. 2e-08;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 EYLKTIKNSLSTWSPSCSVT 21
DB 15 QYLKTIKNSLSTWSPSCSVT 34

RESULT 12
O9TW00 PRELIMINARY; PRT; 80 AA.
AC O9TW00;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota: Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-M0, M7, AND M1;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269955; CAB64238.1; -.
DR EMBL; AJ269958; CAB64238.1; -.
DR EMBL; AJ269959; CAB64239.1; -.
DR EMBL; AJ269956; CAB64241.1; -.
DR InterPro: IPR003067; Circsmpztoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSPL; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9074 MW; A8F40C90DB1C033E CRC64;

Query Match 88.4%; Score 99; DB 5; Length 80;
Best Local Similarity 90.0%; Pred. No. 2e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

Query Match	88.4%;	Score 99;	DB 5;	Length 115
SO SEQUENCE	115 AA;	12926 MM;	697DF630E50B3A90	CRC64;
FT NON-TER	1	1		

Search completed: December 6, 2002, 12:45:15
Job time : 24.0984 secs

THIS PAGE BLANK (USP)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:38:35 : Search time 28.8525 Seconds
(without alignments)
92.367 Million cell updates/sec

Title: US-09-931-325C-148
Perfect score: 108
Sequence: 1 EYLNKIQNSLSTEMSPCSVT 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A.Geneseq_101002.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	100.0	20	19 AAM29734	Universal malarial
2	108	100.0	20	19 AAM61554	T-cell epitope 2.
3	108	100.0	20	23 AA093860	P. falciparum MSP1
4	108	100.0	20	23 AA093924	Hepatitis B virus
5	108	100.0	20	23 AA087777	Malarial epitope s
6	108	100.0	21	23 AA093880	P. falciparum univ
7	108	100.0	21	23 AA087735	Malarial universal
8	108	100.0	22	23 AA087718	Malarial universal
9	108	100.0	26	23 AA093980	Modified portion o
10	108	100.0	26	23 AA087793	Modified C-terminu

11	108	100.0	40	16 AAR87213	P. falciparum deriv
12	108	100.0	171	23 AA093972	Immunogenic HBC ch
13	108	100.0	191	23 AA093974	Chimeric Hepatitis
14	108	100.0	195	23 AA093975	Chimeric Hepatitis
15	108	100.0	424	14 AAR37797	RTS* protein. Syn
16	99	91.7	20	23 AA093803	P. falciparum CS B
17	99	91.7	20	23 AA087778	Malarial epitope s
18	99	91.7	27	23 AA093981	Modified portion o
19	99	91.7	27	23 AA087794	Modified C-terminu
20	93	86.1	20	22 AAM99548	Vaccine related MH
21	93	86.1	54	16 AAR71644	Circumsporozoite R
22	93	86.1	180	11 AAR07290	Circumsporozoite a
23	93	86.1	184	11 AAR07289	Circumsporozoite a
24	93	86.1	250	7 AAP60412	Synthetic peptide
25	93	86.1	309	12 AAR3175	NS1.81-RLfauth.
26	93	86.1	319	11 AAR07945	NS181RLfauth plas
27	93	86.1	319	12 AAR3176	NS1.81-RLfauth. P
28	93	86.1	327	12 AAR3177	NS1.81-RLfauth + (
29	93	86.1	335	12 AAR3178	NS1.81(NAP)ARLfa
30	93	86.1	335	12 AAR3179	NS1.81(NAP)4RLfa
31	93	86.1	411	9 AAP83144	Sequence encoded b
32	93	86.1	412	7 AAP60416	CS protein of mala
33	93	86.1	412	9 AAP80835	Sequence encoded b
34	93	86.1	424	14 AAR37796	RTS protein. Syn
35	82	75.9	402	8 AAP70709	Plasmodium cynomol
36	79	73.1	161	11 AAR07287	Circumsporozoite a
37	77	71.3	21	23 AA093887	P. vivax universal
38	77	71.3	21	23 AA087742	Malarial universal
39	77	71.3	160	11 AAR07288	Circumsporozoite a
40	77	71.3	378	14 AAR30609	Plasmodium vivax c
41	77	71.3	429	10 AAP90064	Antigenic protein
42	77	71.3	1807	22 AAB85687	Recombinant protei
43	77	71.3	2028	22 AAB85688	Recombinant protei
44	75	69.4	378	8 AAP70708	Sequence encoding
45	72	66.7	19	23 AA093861	Plasmodium vivax T

ALIGNMENTS

RESULT 1
AAM29734
ID AAM29734 standard; peptide: 20 AA.
XX
AC AAM29734:
XX
DT 26-OCF-1998 (first entry)
XX
XX Universal malarial T-cell epitope.
DE
XX
XX Circumsporozoite protein; CS; CD4+ T-cell clone; sporozoite; cytotoxic;
KW non-cytotoxic; Class II-restricted human CD4+ T-cell clone; polyclonal;
KW class-I-restricted CD8+ CTL clone; B-cell epitope; vaccine; malaria.
XX
XX Plasmodium falciparum.
OS
XX
XX W09830237-A1.
PN
XX
XX 16-JUL-1998.
PD
XX
XX 24-DEC-1997: 97WO-US24283.
PF
XX
XX 24-DEC-1996: 96US-0034506.
PR
XX
XX (UYNY) UNIV NEW YORK STATE.
PA
XX
XX Nardin E. Nussenzweig RS, Rose K;
PI WPI: 1998-398801/34.
XX
XX
XX Immunogenic composition containing T cell epitope of malaria-derived
PT peptide - incorporated into a polyoxime and optionally B cell
PT epitope, used in vaccines to protect against malaria in subjects of

PT differing genetic background(s)
 XX
 PS Disclosure: p3; 36pp; English.
 CC
 CC The present sequence is a peptide comprising an universal malarial T-cell
 CC epitope, which is contained in the COOH-terminus, amino acid residues
 CC 326-345 of the Plasmodium falciparum NF54 strain circumsporozoite (CS)
 CC protein. This peptide as well as another T-cell epitope were identified
 CC by the use of CD4+ T-cell clones derived from sporozoite immunised
 CC individuals. It is this epitope that has been shown to be recognised by
 CC cytotoxic and non-cytotoxic class II-restricted human CD4+ T-cell clones
 CC and class-I-restricted CD8+ CTL clones. The epitope can be incorporated
 CC into a polyoxime with B-cell epitopes to produce a vaccine that is
 CC protective against malaria in individuals of different genetic
 CC backgrounds.
 XX
 SQ Sequence 20 AA:
 Query Match 100.0%; Score 108; DB 19; Length 20;
 Best Local Similarity 100.0%; Pred. No. 5.9e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 EYLKIQNSLSTWSPSCSVT 20
 1 EYLKIQNSLSTWSPSCSVT 20
 DB
 RESULT 2
 AAM61554
 ID AAM61554 standard; peptide: 20 AA.
 XX
 AC AAM61554;
 XX
 DT 19-OCR-1998 (first entry)
 XX
 DE T-cell epitope 2.
 XX
 KM T-cell; malaria; immunogenic; anti-malarial; prophylactic immunity.
 XX
 OS Plasmodium falciparum
 XX
 PN WO9831382-A1.
 XX
 PD 23-JUL-1998.
 XX
 PF 21-JAN-1998; 98WO-US01527.
 XX
 PR 21-JAN-1997; 97US-0033916.
 (UYNV) UNIV NEW YORK STATE.
 PI Moreno A, Nardin E;
 XX
 DR WPI: 1998-413810/35.
 XX
 PT New immunogenic compositions for malaria - comprise malaria derived
 PT peptide comprising universal T-cell epitope which elicits
 PT anti-malarial T-cell response
 XX
 PS Claim 5; Page 25; 38pp; English.
 XX
 CC The T-cell epitope derived from malaria can be used in an immunogenic
 CC composition. The T-cell epitope elicits an anti-malarial T-cell response
 CC in mammals of diverse genetic backgrounds. The composition can be used
 CC as a vaccine to confer prophylactic or therapeutic immunity against
 CC malaria. They may also be used to inhibit the propagation of a malarial
 CC organism in a susceptible animal.
 CC
 XX
 SQ Sequence 20 AA:
 Query Match 100.0%; Score 108; DB 19; Length 20;
 Best Local Similarity 100.0%; Pred. No. 5.9e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYLKIQNSLSTWSPSCSVT 20
 1 EYLKIQNSLSTWSPSCSVT 20
 DB
 RESULT 3
 AAU93860
 ID AAU93860 standard; peptide: 20 AA.
 XX
 AC AAU93860;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE P. falciparum MSP1 T cell epitope #4.
 XX
 KM Immunogenic; hepatitis B core; HBC;
 KM vaccine; B cell epitope; T cell epitope; immunostimulant.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO200214478-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 16-AUG-2001; 2001WO-US41759.
 XX
 PR 16-AUG-2000; 2000US-225843P.
 PR 22-AUG-2000; 2000US-226867P.
 PR 15-AUG-2001; 2001US-0930915.
 XX
 PA (APOV-) APOVIA INC.
 XX
 PI Birkett AJ;
 XX
 DR WPI: 2002-257601/30.
 XX
 PT Novel recombinant hepatitis B core protein, displaying one or
 PT hepatitis B core protein, displays immunogenic epitopes at N-terminus,
 PT HBC immunogenic loop with linker for conjugated epitope and C-terminus
 PT
 XX
 PS Disclosure: Page 43; 289pp; English.
 XX
 CC The invention relates to a recombinant hepatitis B core protein,
 CC i.e. a chimeric hepatitis B core (HBC) protein (I), displaying one or
 CC more immunogenic epitopes at the N-terminus, HBC immunogenic loop (L) or
 CC C-terminus, or having a heterologous linker for a conjugated epitope in
 CC (L), and containing a Cys residue at, or near, the C-terminus that
 CC confers enhanced stability to the particles. A vaccine comprising (I) is
 CC useful for inducing an immune response in an inoculated host animal, by
 CC inoculating a host animal with the vaccine, and maintaining that
 CC inoculated animal for a time period sufficient for that animal to
 CC develop an immune response. The immunogenic particles formed using (I)
 CC are substantially free of binding to nucleic acids, and are most stable
 CC than the particle formed from otherwise identical HBC chimera that lacks
 CC the C-terminal residue or in which a C-terminal Cys is replaced by
 CC another residue. The chimera particles are most stable on storage in
 CC aqueous compositions that are particles of similar sequence that lack any
 CC C-terminal Cys residues. The chimera molecule exhibits the self-assembly
 CC not exhibiting the nucleic acid binding of those native particles, and
 CC excellent B cell and T cell immunogenicities. The chimera particles are
 CC typically prepared in higher yield than similar particles that are free
 CC of a C-terminal Cys. The particles are often far more immunogenic than
 CC the similar conjugates that lack a C-terminal Cys. Immunogenicities of
 CC particles assembled from the chimera molecules are enhanced as compared to
 CC similar particles assembled from chimera molecules lacking at least one
 CC C-terminal Cys. AAU93802-AAU9397 represent immunogenic HBC particles
 CC amino acid sequences and related sequences of the invention.
 XX
 SQ Sequence 20 AA:
 Query Match 100.0%; Score 108; DB 23; Length 20;

Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYLNKIONSISTEMSPCSVT 20
|||||
Db 1 EYLNKIONSISTEMSPCSVT 20

RESULT 4
AAU93924

ID AAU93924 standard; Peptide; 20 AA.

AC AAU93924;

DT 02-JUL-2002 (first entry)

DE Hepatitis B virus HBC149 peptide #14.

XX Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; HBC;
KW vaccine; B cell epitope; T cell epitope; immunostimulant.

OS Hepatitis B virus.

PN WO200214478-A2.

PD 21-FEB-2002.

PF 16-AUG-2001; 2001WO-US41759.

PR 16-AUG-2000; 2000US-225843P.

PR 22-AUG-2000; 2000US-226867P.

PR 15-AUG-2001; 2001US-0930915.

PA (APOV-) APOVIA INC.

PI Birckett AJ;

DR WPI; 2002-257601/30.

XX Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric
PT hepatitis B core protein, displays immunogenic epitopes at N-terminus,
PT Hbc immunogenic loop with linker for conjugated epitope and C-terminus
PT -
PS
XX

Example 8; Page 130; 289pp; English.

CC The invention relates to a recombinant hepadnavirus nucleocapsid protein,
CC i.e. a chimeric hepatitis B core (Hbc) protein (I), displaying one or
CC more immunogenic epitopes at the N-terminus, Hbc immunogenic loop (L) or
CC C-terminus, or having a heterologous linker for a conjugated epitope in
CC (L), and containing a Cys residue at, or near, the C-terminus that
CC confers enhanced stability to the particles. A vaccine comprising (I) is
CC useful for inducing an immune response in an inoculated host animal, by
CC inoculating a host animal with the vaccine, and maintaining that
CC inoculated animal for a time period sufficient for that animal to
CC develop an immune response. The immunogenic particles formed using (I)
CC are substantially free of binding to nucleic acids, and are most stable
CC than the particle formed from otherwise identical Hbc chimera that lacks
CC the C-terminal residue or in which a C-terminal Cys is replaced by
CC another residue. The chimera particles are most stable on storage in
CC aqueous compositions that are particles of similar sequence that lack any
CC C-terminal Cys residues. The chimera molecule exhibits the self-assembly
CC not exhibiting the nucleic acid binding of those native particles, and
CC excellent B cell and T cell immunogenicities. The chimera particles are
CC typically prepared in higher yield than similar particles that are free
CC of a C-terminal Cys. The particles are often far more immunogenic than
CC the similar conjugates that lack a C-terminal Cys. Immunogenicities of
CC particles assembled from the chimera molecules are enhanced as compared to
CC similar particles assembled from chimera molecules lacking at least one
CC C-terminal Cys. AAU93802-AAU93997 represent immunogenic Hbc particles
CC amino acid sequences and related sequences of the invention.
XX
SQ

Sequence 20 AA;

Query Match 100.0%; Score 108; DB 23; Length 20;

Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYLNKIONSISTEMSPCSVT 20
|||||
Db 1 EYLNKIONSISTEMSPCSVT 20

RESULT 5
AAU87777

ID AAU87777 standard; Peptide; 20 AA.

AC AAU87777;

DT 21-MAY-2002 (first entry)

DE Malarial epitope sequence #4.

XX Hepatitis B virus; nucleocapsid protein; Hbc; hepatitis B virus core;
KW B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoa;
KW circumsporozoite; human immunodeficiency virus type 1; human; squirrel;
KW woodchuck.

OS Plasmodium falciparum.

PN WO200213765-A2.

PD 21-FEB-2002.

PF 16-AUG-2001; 2001WO-US25625.

PR 16-AUG-2000; 2000US-225843P.

PR 15-AUG-2001; 2001US-0931325.

PA (APOV-) APOVIA INC.

PI Birckett AJ;

DR WPI; 2002-241832/29.

XX Recombinant hepatitis B virus core protein chimera molecule, useful to
PT induce antibodies to malarial parasites, comprises malaria-specific
PT T-cell epitope and is engineered for enhanced stability -
PS
XX

Claim 25; Page 110; 197pp; English.

CC The invention relates to a recombinant hepatitis B virus core (Hbc)
CC protein chimera molecule that contains 4 peptide-linked amino acid residue
CC sequence domains. The molecule of the invention contains a region
CC constituting a B cell epitope of the circumsporozoite protein of a
CC species of the parasite, Plasmodium. The chimera sequence is useful as an
CC immunogen for inducing antibodies to the malaria-causing parasite,
CC Plasmodium, particularly P. falciparum and P. vivax. Sequences
CC AAU87695-AAU87804 represent peptide epitopes of the invention.
XX
SQ

Sequence 20 AA;

Query Match 100.0%; Score 108; DB 23; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYLNKIONSISTEMSPCSVT 20
|||||
Db 1 EYLNKIONSISTEMSPCSVT 20

RESULT 6
AAU93880

ID AAU93880 standard; Peptide; 21 AA.

AC AAU93880;

```

XX 02-JUL-2002 (first entry)
DT
XX
XX P. falciparum universal T cell epitope Pf-UTC.
DE
XX
XX Immunogenic: hepadnavirus nucleocapsid protein: hepatitis B core; HBC;
KM vaccine; B cell epitope; T cell epitope; immunostimulant.
OS
XX Plasmodium falciparum.
XX
XX WO200214478-A2.
XX
XX 21-FEB-2002.
PD
XX
XX 16-AUG-2001; 2001WO-US41759.
XX
XX 16-AUG-2000; 2000US-225843P.
XX
XX 22-AUG-2000; 2000US-226867P.
XX
XX 15-AUG-2001; 2001US-0930915.
XX
XX (APOV-) APOVIA INC.
XX
XX Birkett AJ;
XX
XX WPI; 2002-257601/30.
XX
XX
XX Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric
XX hepatitis B core protein, displays immunogenic epitopes at N-terminus,
XX HBC immunogenic loop with linker for conjugated epitope and C-terminus
XX
XX
XX Example 1; Page 110; 289pp; English.
XX
XX The invention relates to a recombinant hepadnavirus nucleocapsid protein,
XX i.e. a chimeric hepatitis B core (HBC) protein (1), displaying one or
XX more immunogenic epitopes at the N-terminus, HBC immunogenic loop (L) or
XX C-terminus, or having a heterologous linker for a conjugated epitope in
XX (L), and containing a Cys residue at, or near, the C-terminus that
XX confers enhanced stability to the particles. A vaccine comprising (1) is
XX useful for inducing an immune response in an inoculated host animal, by
XX inoculating a host animal with the vaccine, and maintaining that
XX inoculated animal for a time period sufficient for that animal to
XX develop an immune response. The immunogenic particles formed using (1)
XX are substantially free of binding to nucleic acids, and are most stable
XX than the particle formed from otherwise identical HBC chimera that lacks
XX the C-terminal residue or in which a C-terminal Cys is replaced by
XX another residue. The chimera particles are most stable on storage in
XX aqueous compositions that are particles of similar sequence that lack any
XX C-terminal Cys residues. The chimera molecule exhibits the self-assembly
XX not exhibiting the nucleic acid binding of those native particles, and
XX excellent B cell and T cell immunogenicities. The chimera particles are
XX typically prepared in higher yield than similar particles that are free
XX of a C-terminal Cys. The particles are often far more immunogenic than
XX the similar conjugates that lack a C-terminal Cys. Immunogenicities of
XX particles assembled from the chimera molecules are enhanced as compared to
XX similar particles assembled from chimera molecules lacking at least one
XX C-terminal Cys. AA093802-AA093997 represent immunogenic HBC particles
XX amino acid sequences and related sequences of the invention.
XX
XX
XX Sequence 21 AA:
XX
XX Query Match 100.0%; Score 108; DB 23; Length 21;
XX Best Local Similarity 100.0%; Pred. No. 6.3e-10;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 EYLKNIQNSTEWSPCSVT 20
XX | | | | | | | | | | | | | | | | | | | | | |
XX Db 2 EYLKNIQNSTEWSPCSVT 21
XX
XX RESULT 7
XX AA087735
XX ID AA087735 standard; Peptide: 21 AA.

```

```

XX
XX AA087735;
XX
XX 21-MAY-2002 (first entry)
XX
XX
XX Malarial universal T cell epitope #4.
DE
XX
XX Hepatitis B virus; nucleocapsid protein; HBC; hepatitis B virus core;
XX B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoacide;
XX circumsporozoite; human immunodeficiency virus type I; human; squirrel;
XX woodchuck.
XX
XX Plasmodium falciparum.
XX
XX WO200213765-A2.
XX
XX 21-FEB-2002.
XX
XX 16-AUG-2001; 2001WO-US25625.
XX
XX 16-AUG-2000; 2000US-225813P.
XX
XX 15-AUG-2001; 2001US-0931325.
XX
XX (APOV-) APOVIA INC.
XX
XX Birkett AJ;
XX
XX WPI; 2002-241832/29.
XX
XX N-PSDB: ABR44248.
XX
XX
XX Recombinant hepatitis B virus core protein chimera molecule, useful to
XX induce antibodies to malarial parasites, comprises malaria-specific
XX T-cell epitope and is engineered for enhanced stability -
XX
XX Example 1; Fig 2B; 197pp; English.
XX
XX The invention relates to a recombinant hepatitis B virus core (HBC)
XX protein chimera molecule that contains 4 peptide-linked amino acid residue
XX sequence domains. The molecule of the invention contains a region
XX constituting a B cell epitope of the circumsporozoite protein of a
XX species of the parasite, Plasmodium. The chimera sequence is useful as an
XX immunogen for inducing antibodies to the malaria-causing parasite,
XX Plasmodium, particularly P. falciparum and P. vivax. Sequences
XX AA087695-AA087804 represent peptide epitopes of the invention.
XX
XX
XX Sequence 21 AA:
XX
XX Query Match 100.0%; Score 108; DB 23; Length 21;
XX Best Local Similarity 100.0%; Pred. No. 6.3e-10;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 EYLKNIQNSTEWSPCSVT 20
XX | | | | | | | | | | | | | | | | | | | | | |
XX Db 2 EYLKNIQNSTEWSPCSVT 21
XX
XX RESULT 8
XX AA087718
XX ID AA087718 standard; Peptide: 22 AA.
XX
XX AA087718;
XX
XX 21-MAY-2002 (first entry)
XX
XX
XX Malarial universal T cell epitope #1.
XX
XX Hepatitis B virus; nucleocapsid protein; HBC; hepatitis B virus core;
XX B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoacide;
XX circumsporozoite; human immunodeficiency virus type I; human; squirrel;
XX woodchuck.
XX
XX Plasmodium falciparum.
XX
XX

```

PM WO200213765-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 16-AUG-2001; 2001WO-US25625.
 XX
 PR 16-AUG-2000; 2000US-225813P.
 PR 15-AUG-2001; 2001US-0931325.
 XX
 PA (APOV-) APOVIA INC.
 XX
 PI Birkett AJ;
 XX
 DR WPI; 2002-241832/29.
 XX
 PT Recombinant hepatitis B virus core protein chimera molecule, useful to
 PT induce antibodies to malarial parasites, comprises malaria-specific
 PT T-cell epitope and is engineered for enhanced stability -
 XX
 XX Claim 51; Page 38; 197pp; English.
 CC The invention relates to a recombinant hepatitis B virus core (HBC)
 CC protein chimera molecule that contains 4 peptide-linked amino acid residue
 CC sequence domains. The molecule of the invention contains a region
 CC constituting a B cell epitope of the circumsporozoite protein of a
 CC species of the parasite, Plasmodium. The chimera sequence is useful as an
 CC immunogen for inducing antibodies to the malaria-causing parasite,
 CC Plasmodium, particularly P. falciparum and P. vivax. Sequences
 CC AA087695-AA087804 represent peptide epitopes of the invention.
 XX
 SQ Sequence 22 AA;
 XX
 Query Match 100.0%; Score 108; DB 23; Length 22;
 Best Local Similarity 100.0%; Pred. No. 6,6e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EYLKNIQNSLSTEWSPSCSVT 20
 DB 3 EYLKNIQNSLSTEWSPSCSVT 22
 XX
 RESULT 9
 AA093980
 ID AA093980 standard; Peptide; 26 AA.
 XX
 AC AA093980;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Modified portion of Hepatitis B core.
 XX
 DE Immunogenic; hepatitis virus nucleocapsid protein; hepatitis B core; HBC;
 KM vaccine; B cell epitope; T cell epitope; immunostimulant.
 XX
 OS Hepatitis B virus.
 OS
 PN WO200214478-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 16-AUG-2001; 2001WO-US41759.
 XX
 PR 16-AUG-2000; 2000US-225843P.
 PR 22-AUG-2000; 2000US-226867P.
 PR 15-AUG-2001; 2001US-0930915.
 XX
 PA (APOV-) APOVIA INC.
 XX
 PI Birkett AJ;
 XX
 DR WPI; 2002-257601/30.
 XX
 PT Novel recombinant hepatitis virus nucleocapsid protein, termed as chimeric

PT hepatitis B core protein, displays immunogenic epitopes at N-terminus,
 PT HBC immunogenic loop with linker for conjugated epitope and C-terminus
 XX
 PS Example 22; Page 163; 289pp; English.
 XX
 XX The invention relates to a recombinant hepatitis virus nucleocapsid protein,
 CC i.e. a chimeric hepatitis B core (HBC) protein (I), displaying one or
 CC more immunogenic epitopes at the N-terminus, HBC immunogenic loop (L) or
 CC C-terminus, or having a heterologous linker for a conjugated epitope in
 CC (L), and containing a Cys residue at, or near, the C-terminus that
 CC confers enhanced stability to the particles. A vaccine comprising (I) is
 CC useful for inducing an immune response in an inoculated host animal, by
 CC inoculating a host animal with the vaccine, and maintaining that
 CC inoculated animal for a time period sufficient for that animal to
 CC develop an immune response. The immunogenic particles formed using (I)
 CC are substantially free of binding to nucleic acids, and are most stable
 CC than the particle formed from otherwise identical HBC chimera that lacks
 CC the C-terminal residue or in which a C-terminal Cys is replaced by
 CC another residue. The chimera particles are most stable on storage in
 CC aqueous compositions that are particles of similar sequence that lack any
 CC C-terminal Cys residues. The chimera molecule exhibits the self-assembly
 CC not exhibiting the nucleic acid binding of those native particles, and
 CC excellent B cell and T cell immunogenicities. The chimera particles are
 CC typically prepared in higher yield than similar particles that are free
 CC of a C-terminal Cys. The particles are often far more immunogenic than
 CC the similar conjugates that lack a C-terminal Cys. Immunogenicities of
 CC particles assembled from the chimera molecules are enhanced as compared to
 CC similar particles assembled from chimera molecules lacking at least one
 CC C-terminal Cys. AA093802-AA093997 represent immunogenic HBC particles
 CC amino acid sequences and related sequences of the invention.
 XX
 SQ Sequence 26 AA;
 XX
 Query Match 100.0%; Score 108; DB 23; Length 26;
 Best Local Similarity 100.0%; Pred. No. 8e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EYLKNIQNSLSTEWSPSCSVT 20
 DB 7 EYLKNIQNSLSTEWSPSCSVT 26
 XX
 RESULT 10
 AA087793
 ID AA087793 standard; Peptide; 26 AA.
 XX
 AC AA087793;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Modified C-terminus of Hepatitis B virus nucleocapsid protein #13.
 XX
 DE Hepatitis B virus; nucleocapsid protein; HBC; hepatitis B virus core;
 KM B cell epitope; T cell epitope; malaria; HIV-1; Plasmodium; protozoacide;
 KM circumsporozoite; human immunodeficiency virus type I; human; squirrel;
 KM woodchuck.
 XX
 OS Hepatitis B virus.
 OS Synthetic.
 OS
 PN WO200213765-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 16-AUG-2001; 2001WO-US25625.
 XX
 PR 16-AUG-2000; 2000US-225813P.
 PR 15-AUG-2001; 2001US-0931325.
 XX
 PA (APOV-) APOVIA INC.
 XX
 PI Birkett AJ;

XX	WP1: 2002-241832/29.
DR	
XX	Recombinant hepatitis B virus core protein chimera molecule, useful to
PT	induce antibodies to malarial parasites, comprises malaria-specific
PT	T-cell epitope and is engineered for enhanced stability -
XX	
PS	Example 13; Page 128; 197pp; English.
XX	
CC	The invention relates to a recombinant hepatitis B virus core (HBC)
CC	protein chimera molecule that contains 4 peptide-linked amino acid residue
CC	sequence domains. The molecule of the invention contains a region
CC	constituting a B cell epitope of the circumsporozoite protein of a
CC	species of the parasite, Plasmodium. The chimera sequence is useful as an
CC	immunogen for inducing antibodies to the malaria-causing parasite,
CC	Plasmodium, particularly P. falciparum and P. vivax. Sequences
CC	AA087695-AA087804 represent peptide epitopes of the invention.
XX	
SQ	Sequence 26 AA:
XX	
Query Match	100.0%; Score 108; DB 23; Length 26;
Best Local Similarity	100.0%; Pred. No. 8e-10;
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY	1 EYLKIKNSLSTWSPGCVT 20
DB	7 EYLKIKNSLSTWSPGCVT 26
XX	
RESULT 11	
ID	AA087213
XX	AA087213 standard; peptide: 40 AA.
AC	
XX	
XX	AA087213:
DT	16-MAY-1996 (first entry)
XX	
DE	P. falciparum derived presented KEKE-like motif contg. peptide.
XX	
KW	Proteasome; activation; cell-mediated immunity; immunogen; tolerance;
XX	KEKE motif; interleukin.
XX	
OS	Plasmodium falciparum.
XX	
EH	Key
EH	Location/Qualifiers
FT	1..20
FT	/note= "KEKE-like sequence"
FT	Misc-difference 21
FT	/note= "undefined linker sequence of 34
	amino acids"
FT	22..40
FT	/note= "presented peptide"
XX	
XX	Peptide
XX	
XX	WO9527058-A1.
XX	
XX	12-OCT-1995.
XX	
XX	01-APR-1994; 94WO-US03591.
XX	
XX	01-APR-1994; 94WO-US03591.
XX	
PA	(UTAH) UNIV UTAH.
XX	
PI	Realini CA, Rechsteiner MC;
XX	
XX	WP1: 1995-358633/46.
XX	
XX	DNA encoding human proteasome activator - used to elicit
XX	cell-mediated immunity or tolerance to a selected immunogenic
XX	peptide
XX	
XX	
XX	Disclosure; Page 45; 71pp; English.
XX	

CC	AAR87210-887215 are, presented peptides contg. a KEKE-like motif.
CC	KEKE motif contg. peptides can be used in a method for inducing cell-
CC	mediated immunity against or tolerance to specific epitopes using
CC	plasmids encoding a human proteasome activator (PA) and appropriate
CC	epitope-bearing peptides adjacent to presentation marker peptides
CC	(lysine and glutamine rich peptides termed KEKE motifs peptides, that
CC	mark adjacent peptides for presentation). The method can be used to
CC	elicit cell- mediated immunity or tolerance to pathogen-encoded peptides
CC	or tumour specific antigens.
XX	
SQ	Sequence 40 AA:
	Query Match 100.0%; Score 108; DB 16; Length 40;
	Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1 EYLKNIQNLSLSTWSPSCSVT 20
Db	1 EYLKNIQNLSLSTWSPSCSVT 20
RESULT 12	
ID	AAU93972
AC	AAU93972 standard; Peptide: 171 AA.
XX	
DT	AAU93972;
XX	
DE	02-JUL-2002 (first entry)
XX	
KM	Immunogenic HBC chimeric particle #9.
XX	
OS	Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; HBC;
XX	vaccine; B cell epitope; T cell epitope; Immunostimulant.
PN	Plasmodium falciparum.
WO200214478-A2.	
PD	
PB	21-FEB-2002.
PF	
PR	16-AUG-2001; 2001WO-US41759.
PR	16-AUG-2000; 2000US-225843P.
PR	22-AUG-2000; 2000US-226867P.
PR	15-AUG-2001; 2001US-0930915.
PA	(APOV-) APOVIA INC.
Birkett AJ;	
WP1: 2002-257601/30.	
Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric	
hepatitis B core protein, displays immunogenic epitopes at N-terminus,	
HBC immunogenic loop with linker for conjugated epitope and C-terminus	
-	
Example 4; Page 273; 289pp; English.	
The invention relates to a recombinant hepadnavirus nucleocapsid protein,	
i.e. a chimeric hepatitis B core (HBC) protein (I), displaying one or	
more immunogenic epitopes at the N-terminus, HBC immunogenic loop (L) or	
C-terminus, or having a heterologous linker for a conjugated epitope in	
(L), and containing a Cys residue at, or near, the C-terminus that	
confers enhanced stability to the particles. A vaccine comprising (I) is	
useful for inducing an immune response in an inoculated host animal, by	
inoculating a host animal with the vaccine, and maintaining that	
inoculated animal for a time period sufficient for that animal to	
develop an immune response. The immunogenic particles formed using (I)	
are substantially free of binding to nucleic acids, and are most stable	
than the particle formed from otherwise identical HBC chimera that lacks	
the C-terminal residue or in which a C-terminal Cys is replaced by	
another residue. The chimera particles are most stable on storage in	

aqueous compositions that are particles of similar sequence that lack any C-terminal Cys residues. The chimera molecule exhibits the self-assembly not exhibiting the nucleic acid binding of those native particles, and excellent B cell and T cell immunogenicities. The chimera particles are typically prepared in higher yield than similar particles that are free of a C-terminal Cys. The particles are often far more immunogenic than the similar conjugates that lack a C-terminal Cys. Immunogenicities of particles assembled from the chimera molecules are enhanced as compared to similar particles assembled from chimera molecules lacking at least one C-terminal Cys. AAU93802-AAU93997 represent immunogenic HBC particles amino acid sequences and related sequences of the invention.

Sequence 171 AA:

Query Match 100.0%; Score 108; DB 23; Length 171;
Best Local Similarity 100.0%; Pred. No. 7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 EYLKNTKNTSLSTWSPSCSVT 20
152 EYLKNTKNTSLSTWSPSCSVT 171

RESULT 13

AAU93974
ID AAU93974 standard; Peptide; 191 AA.

AAU93974;

02-JUL-2002 (first entry)

Chimeric Hepatitis B virus particle V2.pfl.

Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; HBC; vaccine; B cell epitope; T cell epitope; immunostimulant.

Plasmodium falciparum.

WO200214478-A2.

21-FEB-2002.

16-AUG-2001; 2001WO-US41759.

16-AUG-2000; 2000US-225843P.

22-AUG-2000; 2000US-226867P.

15-AUG-2001; 2001US-0930915.

(APOV-) APOVIA INC.

Birkett AJ;

WPI; 2002-257601/30.

Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric hepatitis B core protein, displays immunogenic epitopes at N-terminus, HBC immunogenic loop with linker for conjugated epitope and C-terminus

Example 8; Page 276; 289pp; English.

The invention relates to a recombinant hepadnavirus nucleocapsid protein, i.e. a chimeric hepatitis B core (HBC) protein (I), displaying one or more immunogenic epitopes at the N-terminus, HBC immunogenic loop (L) or C-terminus, or having a heterologous linker for a conjugated epitope in (L), and containing a Cys residue at, or near, the C-terminus that confers enhanced stability to the particles. A vaccine comprising (I) is useful for inducing an immune response in an inoculated host animal, by inoculating a host animal with the vaccine, and maintaining that inoculated animal for a time period sufficient for that animal to develop an immune response. The immunogenic particles formed using (I) are substantially free of binding to nucleic acids, and are most stable than the particle formed from otherwise identical HBC chimera that lacks

the C-terminal residue or in which a C-terminal Cys is replaced by another residue. The chimera particles are most stable on storage in aqueous compositions that are particles of similar sequence that lack any C-terminal Cys residues. The chimera molecule exhibits the self-assembly not exhibiting the nucleic acid binding of those native particles, and excellent B cell and T cell immunogenicities. The chimera particles are typically prepared in higher yield than similar particles that are free of a C-terminal Cys. The particles are often far more immunogenic than the similar conjugates that lack a C-terminal Cys. Immunogenicities of particles assembled from the chimera molecules are enhanced as compared to similar particles assembled from chimera molecules lacking at least one C-terminal Cys. AAU93802-AAU93997 represent immunogenic HBC particles amino acid sequences and related sequences of the invention.

Sequence 191 AA:

Query Match 100.0%; Score 108; DB 23; Length 191;
Best Local Similarity 100.0%; Pred. No. 7.9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 EYLKNTKNTSLSTWSPSCSVT 20
172 EYLKNTKNTSLSTWSPSCSVT 191

RESULT 14

AAU93975
ID AAU93975 standard; Peptide; 195 AA.

AAU93975;

02-JUL-2002 (first entry)

Chimeric Hepatitis B virus particle V17P13.1.

Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; HBC; vaccine; B cell epitope; T cell epitope; immunostimulant.

Plasmodium falciparum.

WO200214478-A2.

21-FEB-2002.

16-AUG-2001; 2001WO-US41759.

16-AUG-2000; 2000US-225843P.

22-AUG-2000; 2000US-226867P.

15-AUG-2001; 2001US-0930915.

(APOV-) APOVIA INC.

Birkett AJ;

WPI; 2002-257601/30.

Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric hepatitis B core protein, displays immunogenic epitopes at N-terminus, HBC immunogenic loop with linker for conjugated epitope and C-terminus

Example 20; Page 277-278; 289pp; English.

The invention relates to a recombinant hepadnavirus nucleocapsid protein, i.e. a chimeric hepatitis B core (HBC) protein (I), displaying one or more immunogenic epitopes at the N-terminus, HBC immunogenic loop (L) or C-terminus, or having a heterologous linker for a conjugated epitope in (L), and containing a Cys residue at, or near, the C-terminus that confers enhanced stability to the particles. A vaccine comprising (I) is useful for inducing an immune response in an inoculated host animal, by inoculating a host animal with the vaccine, and maintaining that inoculated animal for a time period sufficient for that animal to develop an immune response. The immunogenic particles formed using (I)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:42:06 : Search time 11.4754 Seconds
(without alignments)
51.280 Million cell updates/sec

Title: US-09-931-325C-148

Perfect score: 108

Sequence: 1 EYLKIONSLSSTWSPCSVT 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Optimal number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	424	2	US-08-760-797A-3 Sequence 3, Appl1
2	108	100.0	424	4	US-08-932-929B-3 Sequence 3, Appl1
3	93	86.1	412	1	US-08-313-288B-18 Sequence 18, Appl1
4	93	86.1	423	2	US-08-760-797A-1 Sequence 1, Appl1
5	93	86.1	424	4	US-08-932-929B-1 Sequence 1, Appl1
6	77	71.3	20	1	US-07-848-636B-3 Sequence 3, Appl1
7	77	71.3	378	6	5171843-9 Patent No. 5171843
8	60	55.6	478	3	US-08-155-88B-2 Sequence 2, Appl1
9	50	46.3	46	1	US-08-395-602A-5 Sequence 5, Appl1
10	50	46.3	46	2	US-08-021-625D-5 Sequence 2, Appl1
11	50	46.3	525	4	US-09-369-364A-21 Sequence 21, Appl1
12	49	45.4	18	1	US-08-395-602A-1 Sequence 1, Appl1
13	49	45.4	18	2	US-08-021-625D-1 Sequence 1, Appl1
14	49	45.4	23	1	US-08-395-602A-2 Sequence 2, Appl1
15	49	45.4	23	2	US-08-021-625D-2 Sequence 2, Appl1
16	49	45.4	23	4	US-08-986-659B-9 Sequence 9, Appl1
17	49	45.4	23	4	US-08-986-659B-33 Sequence 33, Appl1
18	49	45.4	33	4	US-08-986-659B-34 Sequence 34, Appl1
19	49	45.4	114	1	US-08-309-604-2 Sequence 2, Appl1
20	49	45.4	126	1	US-08-395-602A-4 Sequence 4, Appl1
21	49	45.4	126	2	US-08-021-625D-4 Sequence 4, Appl1
22	49	45.4	559	1	US-08-313-288B-14 Sequence 14, Appl1
23	46	42.6	38	1	US-08-444-005-18 Sequence 18, Appl1
24	46	42.6	469	1	US-08-313-288B-15 Sequence 15, Appl1
25	46	42.6	656	1	US-08-444-005-15 Sequence 15, Appl1
26	46	42.6	656	4	US-09-069-023-28 Sequence 28, Appl1
27	45	41.7	232	6	5171843-7 Patent No. 5171843

28	44	40.7	9	1	US-07-646-531D-1 Sequence 1, Appl1
29	44	40.7	9	1	US-07-646-531D-8 Sequence 8, Appl1
30	44	40.7	9	1	US-07-646-531D-15 Sequence 15, Appl1
31	44	40.7	9	1	US-07-646-531D-16 Sequence 16, Appl1
32	44	40.7	9	2	US-08-488-273-1 Sequence 1, Appl1
33	44	40.7	9	2	US-08-488-273-8 Sequence 8, Appl1
34	44	40.7	9	2	US-08-488-273-13 Sequence 13, Appl1
35	44	40.7	9	2	US-08-488-273-14 Sequence 14, Appl1
36	44	40.7	9	2	US-08-858-971-1 Sequence 1, Appl1
37	44	40.7	9	4	US-09-197-770B-2 Sequence 14, Appl1
38	44	40.7	9	4	US-09-197-770B-14 Sequence 14, Appl1
39	44	40.7	9	6	5426100-1 Patent No. 5426100
40	44	40.7	9	6	5426100-8 Patent No. 5426100
41	44	40.7	227	4	US-09-182-145-15 Sequence 15, Appl1
42	44	40.7	228	4	US-09-182-145-19 Sequence 19, Appl1
43	44	40.7	228	4	US-09-182-145-77 Sequence 77, Appl1
44	44	40.7	229	4	US-09-182-145-76 Sequence 76, Appl1
45	44	40.7	229	4	US-09-182-145-99 Sequence 99, Appl1

ALIGNMENTS

RESULT 1
US-08-760-797A-3
Sequence 3, Application US/08760797A
Patent No. 5928902
GENERAL INFORMATION:
APPLICANT: De Wilde, Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
TITLE OF INVENTION: from Plasmodium and HBsAg
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08760, 797A
FILING DATE: 04-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/442, 612
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-760-797A-3
Query Match 100.0%; Score 108; DB 2; Length 424;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EYLKIONSLSSTWSPCSVT 20

Db 124 EYLKIONSISTEWSPCSVT 143

RESULT 2

US-08-932-929B-3

Sequence 3, Application US/08932929B

Patent No. 6169171

GENERAL INFORMATION:

APPLICANT: De Wilde, Michel

APPLICANT: Cohen, Joseph

TITLE OF INVENTION: Hybrid Protein Between CS

TITLE OF INVENTION: from Plasmodium and HBSAG

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Smithkline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/932,929B

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/760,797

FILING DATE: 04-DEC-1996

APPLICATION NUMBER: 08/442,612

FILING DATE: 17-MAY-1995

ATTORNEY/AGENT INFORMATION:

NAME: Baumeister, Kirk

REGISTRATION NUMBER: 33,833

REFERENCE/DOCKET NUMBER: B45015-1FMC2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5096

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 424 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

08-932-929B-3

Query Match 100.0%; Score 108; DB 4; Length 424;

Best Local Similarity 100.0%; Pred. No. 1.9e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYLKIONSISTEWSPCSVT 20

Db 124 EYLKIONSISTEWSPCSVT 143

RESULT 3

US-08-313-288B-18

Sequence 18, Application US/08313288B

Patent No. 5750502

GENERAL INFORMATION:

APPLICANT: Jessell, Thomas M. and Avihu Klar

TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/313,288B

FILING DATE: January 5, 1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 40028-A-PCT-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0526

TELEX:

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 412 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-313-288B-18

Query Match 86.1%; Score 93; DB 1; Length 412;

Best Local Similarity 80.0%; Pred. No. 3.1e-06;

Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYLKIONSISTEWSPCSVT 20

Db 333 QYLKIRNSISTEWSPCSVT 352

RESULT 4

US-08-760-797A-1

Sequence 1, Application US/08760797A

Patent No. 5928902

GENERAL INFORMATION:

APPLICANT: De Wilde, Michel

APPLICANT: Cohen, Joseph

TITLE OF INVENTION: Hybrid Protein Between CS

TITLE OF INVENTION: from Plasmodium and HBSAG

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Smithkline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/760,797A

FILING DATE: 04-DEC-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/442,612

FILING DATE: 17-MAY-1995

ATTORNEY/AGENT INFORMATION:

NAME: Baumeister, Kirk

REGISTRATION NUMBER: 33,833

REFERENCE/DOCKET NUMBER: B45015-1C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5096

TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-760-797A-1

Query Match 86.1%; Score 93; DB 2; Length 423;
Best Local Similarity 80.0%; Pred. No. 3.2e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYLNKIONSLSTWSPCSVT 20
Db 127 QYLKKIKNSISTWSPCSVT 146

RESULT 5

S-08-932-929B-1
Sequence 1, Application US/08932929B
Patent No. 6169171
GENERAL INFORMATION:
APPLICANT: De Wilde, Michel
TITLE OF INVENTION: Hybrid Protein Between CS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,929B
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/760,797
FILING DATE: 04-DEC-1996
APPLICATION NUMBER: 08/442,612
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1FWC2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-932-929B-1

Query Match 86.1%; Score 93; DB 4; Length 424;
Best Local Similarity 80.0%; Pred. No. 3.2e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYLNKIONSLSTWSPCSVT 20
Db 128 QYLKKIKNSISTWSPCSVT 147

RESULT 6
US-07-848-636B-3

Sequence 3, Application US/07848636B
Patent No. 5599543
GENERAL INFORMATION:
APPLICANT: Hoffman, Stephen L
APPLICANT: Charoenvit, Yupin
TITLE OF INVENTION: A PHARMACEUTICAL COMPOSITION CONTAINING
FOUR AMINO ACID EPITOPE PROTECTIVE AGAINST PLASMODIUM
TITLE OF INVENTION: VIVAX MALARIA
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Naval Medical Research & Development Command
STREET: 8901 Wisconsin Ave Bldg 1, T-12.
CITY: Bethesda
STATE: Maryland
COUNTRY: USA
ZIP: 20889-5606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/848,636B
FILING DATE: 09-MAR-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,551
FILING DATE: 06-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Spevack, A. David
REGISTRATION NUMBER: 24,743
REFERENCE/DOCKET NUMBER: N.C. 72,634
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 295-6759
TELEFAX: (301) 295-1022
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
PUBLICATION INFORMATION:
DOCUMENT NUMBER: US 5,095,093
FILING DATE: 10-MAR-1992
US-07-848-636B-3

Query Match 71.3%; Score 77; DB 1; Length 20;
Best Local Similarity 60.0%; Pred. No. 2.8e-05;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYLNKIONSLSTWSPCSVT 20
Db 1 EYLDKVRATVCTWSPCSVT 20

RESULT 7

5171843-9
Patent No. 5171843
APPLICANT: NUSSENZWEIG, VICTOR
TITLE OF INVENTION: IMMUNOGENIC POLYPEPTIDE AND METHOD FOR
PURIFYING IT
NUMBER OF SEQUENCES: 13
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/175,112

FILING DATE: 30-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 754,645
FILING DATE: 9-JUL-1985
APPLICATION NUMBER: 115,634
FILING DATE: 26-OCT-1987
APPLICATION NUMBER: 649,903
FILING DATE: 12-SEP-1984
SEQ ID NO: 9
LENGTH: 378
5171843-9

Query Match 71.3%, Score 77; DB 6; Length 378;
Best Local Similarity 60.0%; Pred. No. 0.00067;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 EYLNKQNSLSTWSPCSVT 20
DB 300 EYLDKVRATVGTWTPCSVT 319

RESULT 8
US-08-155-888-2
Sequence 2, Application US/08155888
Patent No. 6066623
GENERAL INFORMATION:
APPLICANT: Hoffman, Stephen L.
APPLICANT: Hedstrom, Richard C.
APPLICANT: Sedegrah, Martha
TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE PROTECTIVE
TITLE OF INVENTION: AGAINST MALARIA, METHODS OF PROTECTION AND VECTOR FOR
DELIVERING POLYNUCLEOTIDE VACCINES
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Naval Medical Res. & Dev. Cmd.
STREET: Bldg. 1, T-12 8901 Wisconsin Ave.
CITY: Bethesda
STATE: Maryland
COUNTRY: USA
ZIP: 20889-5606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/155,888
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Spevack, A. David
REGISTRATION NUMBER: 24,743
REFERENCE/DOCKET NUMBER: N.C. 75,851
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 295-6759
TELEFAX: (202) 295-1022
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-155-888-2

Query Match 55.6%, Score 60; DB 3; Length 478;
Best Local Similarity 50.0%; Pred. No. 0.029;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 EYLNKQNSLSTWSPCSVT 20
DB 401 EYVKQISQLTREMSPCSVT 420

RESULT 9
US-08-395-602A-5
Sequence 5, Application US/08395602A
Patent No. 5766899
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
APPLICANT: Ding, Zhi-Ming
TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,602A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-395-602A-5

Query Match 46.3%, Score 50; DB 1; Length 46;
Best Local Similarity 53.3%; Pred. No. 0.71;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 6 IQNSLSTWSPCSVT 20
DB 22 VDPNANPEWSPCSVT 36

RESULT 10
US-08-021-625D-5
Sequence 5, Application US/08021625D
Patent No. 5876851
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
APPLICANT: Ding, Zhi-Ming
TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,625D
; FILING DATE: 16-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSC:410/HYL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-021-625D-5
;
Query Match 46.3%; Score 50; DB 2; Length 46;
Best Local Similarity 53.3%; Pred. No. 0.71;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 IONSLSTEMSPCSVT 20
; : : |||||
Db 22 VDPNANPEMSPCSVT 36

RESULT 11
; Sequence 21, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-R1
;
US-09-369-364A-21
;
Query Match 46.3%; Score 50; DB 4; Length 525;
Best Local Similarity 72.7%; Pred. No. 9.9;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 10 LSTEMSPCSVT 20
; : : |||||
Db 441 LAQEMSPCSVT 451

RESULT 12
; Sequence 1, Application US/08395602A
; Patent No. 5766899
; GENERAL INFORMATION:
; APPLICANT: Kuo, M. Tien
; APPLICANT: Ding, Zhi-Ming
; TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
; TITLE OF INVENTION: Liver Cells
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
```

```
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,602A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSC:410/HYL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-395-602A-1
;
Query Match 45.4%; Score 49; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 EMSPCSVT 20
; : : |||||
Db 1 EMSPCSVT 8

RESULT 13
; Sequence 1, Application US/08021625D
; Patent No. 5976851
; GENERAL INFORMATION:
; APPLICANT: Kuo, M. Tien
; APPLICANT: Ding, Zhi-Ming
; TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
; TITLE OF INVENTION: Liver Cells
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,625D
; FILING DATE: 16-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSC:410/HYL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
```

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-021-625D-1

Query Match 45.4%; Score 49; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 EMSPCSVT 20
|||||||
Db 1 EMSPCSVT 8

RESULT 14
US-08-395-602A-2
Sequence 2, Application US/08395602A
Patent No. 5766899

GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
APPLICANT: Ding, Zhi-Ming
TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into
TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,602A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-395-602A-2

Query Match 45.4%; Score 49; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 EMSPCSVT 20
|||||||
Db 1 EMSPCSVT 8

RESULT 15
US-08-021-625D-2
Sequence 2, Application US/08021625D
Patent No. 5976851

GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
APPLICANT: Ding, Zhi-Ming
TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into

TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,625D
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-021-625D-2

Query Match 45.4%; Score 49; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 EMSPCSVT 20
|||||||
Db 1 EMSPCSVT 8

Search completed: December 6, 2002, 12:46:44
Job time : 11.4754 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:44:06 : Search time 6.22951 Seconds
(without alignments)
52.146 Million cell updates/sec

Title: US-09-931-325C-148
Perfect score: 108
Sequence: 1 EYLKXIONSLSTWSPSCSVT 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues
Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications-AA:*
1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PCIT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/PCITUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	46.3	53	10	US-09-800-729-164
2	50	46.3	54	10	US-09-858-081-6
3	50	46.3	54	10	US-09-858-068-6
4	50	46.3	54	10	US-10-014-070-9
5	50	46.3	525	9	US-09-978-295A-301
6	50	46.3	525	9	US-09-978-697-301
7	50	46.3	525	9	US-09-978-192A-301
8	50	46.3	525	10	US-09-918-171A-21
9	50	46.3	525	12	US-10-052-586-436
10	50	46.3	1745	9	US-09-800-729-89
11	50	46.3	1762	9	US-10-044-807-2
12	50	46.3	1766	10	US-09-764-853-554
13	45	42.6	656	10	US-09-862-027-30
14	45	41.7	9	12	US-10-042-202-11
15	45	41.7	59	10	US-09-864-761-40588
16	45	41.7	465	10	US-09-869-515-12
17	45	41.7	476	10	US-09-869-515-6
18	45	41.7	969	10	US-09-969-515-10
19	45	41.7	980	10	US-09-969-515-4

20	45	41.7	1213	10	US-09-969-515-8	Sequence 8, Appl1
21	45	41.7	1224	10	US-09-969-515-2	Sequence 2, Appl1
22	45	41.7	2150	10	US-09-321-987B-2	Sequence 2, Appl1
23	45	41.7	2155	10	US-09-800-729-155	Sequence 155, App
24	44	40.7	58	9	US-10-010-408-11	Sequence 11, Appl1
25	44	40.7	226	10	US-09-945-676-8	Sequence 8, Appl1
26	44	40.7	227	9	US-10-010-408-13	Sequence 13, Appl1
27	44	40.7	250	9	US-10-010-408-2	Sequence 2, Appl1
28	44	40.7	250	10	US-09-915-582-53	Sequence 53, Appl1
29	44	40.7	250	10	US-09-915-582-69	Sequence 69, Appl1
30	44	40.7	1588	9	US-10-000-512-2	Sequence 2, Appl1
31	43.5	40.3	851	10	US-09-730-989-4	Sequence 4, Appl1
32	43	39.8	63	12	US-10-036-869-30	Sequence 30, Appl1
33	43	39.8	119	10	US-09-839-185-16	Sequence 16, Appl1
34	43	39.8	372	10	US-09-764-903-45	Sequence 45, Appl1
35	43	39.8	491	9	US-10-125-470-10	Sequence 10, Appl1
36	43	39.8	491	9	US-10-125-452-10	Sequence 10, Appl1
37	42	38.9	56	10	US-09-864-761-34159	Sequence 34159, A
38	42	38.9	60	10	US-09-800-729-167	Sequence 167, App
39	42	38.9	285	9	US-09-982-598-272	Sequence 272, App
40	42	38.9	285	9	US-09-989-293A-272	Sequence 272, App
41	42	38.9	285	10	US-09-989-722-272	Sequence 272, App
42	42	38.9	285	10	US-09-989-723-272	Sequence 272, App
43	42	38.9	285	10	US-09-989-729-272	Sequence 272, App
44	42	38.9	285	10	US-09-989-727-272	Sequence 272, App
45	42	38.9	285	10	US-09-989-731-272	Sequence 272, App

ALIGNMENTS

RESULT 1
US-09-800-729-164
; Sequence 164, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044PI
; CURRENT APPLICATION NUMBER: US/09/800, 729
; PRIOR APPLICATION NUMBER: 2001-03-08
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155, 709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 164
; LENGTH: 53
; TYPE: PRT
; ORGNISM: Homo sapiens
US-09-800-729-164

Query Match 46.3%; Score 50; DB 10; Length 53;
Best Local Similarity 72.7%; Pred. No. 0.28;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 10 LSTWSPSCSVT 20
I: |||||I
Db 2 LAQWSPCTVT 12

RESULT 2
US-09-858-081-6
; Sequence 6, Application US/09858081
; Patent No. US20020072490A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Cook, William James
; APPLICANT: SliOS-Santiago, Immaculada
; TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEINASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-049001

```
; CURRENT APPLICATION NUMBER: US/09/858,081
; CURRENT FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/204,159
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: 60/204,160
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-09-858-081-6
```

```
Query Match      46.3%; Score 50; DB 10; Length 54;
Best Local Similarity 88.9%; Pred. NO. 0.28;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Y      12 TEMSPCSVT 20
      :|||||
Db      4 SEMSPCSVT 12
```

```
RESULT 3
US-09-858-068-6
; Sequence 6, Application US/09858068
; Patent No. US20020076778A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Liebertmann, Rosana
; APPLICANT: Cook, William James
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-057001
; CURRENT APPLICATION NUMBER: US/09/858,068
; CURRENT FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/204,159
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: 60/204,160
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-09-858-068-6
```

```
Query Match      46.3%; Score 50; DB 10; Length 54;
Best Local Similarity 88.9%; Pred. NO. 0.28;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY     12 TEMSPCSVT 20
      :|||||
Db      4 SEMSPCSVT 12
```

```
RESULT 4
US-10-014-070-9
; Sequence 9, Application US/10014070
; Patent No. US20020119555A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhkar
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Spurling, Heidi Lynn
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 53014, A Human Metalloprotease Family
; TITLE OF INVENTION: Member and Uses Therefor
; FILE REFERENCE: MPI2000-523PIRCP1(M)
```

```
; CURRENT APPLICATION NUMBER: US/10/014,070
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/258,373
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus
US-10-014-070-9
```

```
Query Match      46.3%; Score 50; DB 12; Length 54;
Best Local Similarity 88.9%; Pred. NO. 0.28;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY     12 TEMSPCSVT 20
      :|||||
Db      4 SEMSPCSVT 12
```

```
RESULT 5
US-09-978-295A-301
; Sequence 301, Application US/0978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630PIC11
; CURRENT APPLICATION NUMBER: US/09/978,295A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
```

1	PRIOR FILING DATE:	1998-03-11
2	PRIOR APPLICATION NUMBER:	60/077641
3	PRIOR FILING DATE:	1998-03-11
4	PRIOR APPLICATION NUMBER:	60/077649
5	PRIOR FILING DATE:	1998-03-11
6	PRIOR APPLICATION NUMBER:	60/077791
7	PRIOR FILING DATE:	1998-03-12
8	PRIOR APPLICATION NUMBER:	60/078004
9	PRIOR FILING DATE:	1998-03-13
10	PRIOR APPLICATION NUMBER:	60/078886
11	PRIOR FILING DATE:	1998-03-20
12	PRIOR APPLICATION NUMBER:	60/078936
13	PRIOR FILING DATE:	1998-03-20
14	PRIOR APPLICATION NUMBER:	60/078910
15	PRIOR FILING DATE:	1998-03-20
16	PRIOR APPLICATION NUMBER:	60/078939
17	PRIOR FILING DATE:	1998-03-20
18	PRIOR APPLICATION NUMBER:	60/079294
19	PRIOR FILING DATE:	1998-03-25
20	PRIOR APPLICATION NUMBER:	60/079656
21	PRIOR FILING DATE:	1998-03-26
22	PRIOR APPLICATION NUMBER:	60/079664
23	PRIOR FILING DATE:	1998-03-27
24	PRIOR APPLICATION NUMBER:	60/079669
25	PRIOR FILING DATE:	1998-03-27
26	PRIOR APPLICATION NUMBER:	60/079663
27	PRIOR FILING DATE:	1998-03-27
28	PRIOR APPLICATION NUMBER:	60/079728
29	PRIOR FILING DATE:	1998-03-27
30	PRIOR APPLICATION NUMBER:	60/079786
31	PRIOR FILING DATE:	1998-03-27
32	PRIOR APPLICATION NUMBER:	60/079920
33	PRIOR FILING DATE:	1998-03-30
34	PRIOR APPLICATION NUMBER:	60/079923
35	PRIOR FILING DATE:	1998-03-30
36	PRIOR APPLICATION NUMBER:	60/080105
37	PRIOR FILING DATE:	1998-03-31
38	PRIOR APPLICATION NUMBER:	60/080107
39	PRIOR FILING DATE:	1998-03-31
40	PRIOR APPLICATION NUMBER:	60/080165
41	PRIOR FILING DATE:	1998-03-31
42	PRIOR APPLICATION NUMBER:	60/080194
43	PRIOR FILING DATE:	1998-03-31
44	PRIOR APPLICATION NUMBER:	60/080327
45	PRIOR FILING DATE:	1998-04-01
46	PRIOR APPLICATION NUMBER:	60/080328
47	PRIOR FILING DATE:	1998-04-01
48	PRIOR APPLICATION NUMBER:	60/080333
49	PRIOR FILING DATE:	1998-04-01
50	PRIOR APPLICATION NUMBER:	60/080334
51	PRIOR FILING DATE:	1998-04-01
52	PRIOR APPLICATION NUMBER:	60/081070
53	PRIOR FILING DATE:	1998-04-08
54	PRIOR APPLICATION NUMBER:	60/081049
55	PRIOR FILING DATE:	1998-04-08
56	PRIOR APPLICATION NUMBER:	60/081071
57	PRIOR FILING DATE:	1998-04-08
58	PRIOR APPLICATION NUMBER:	60/081195
59	PRIOR FILING DATE:	1998-04-08
60	PRIOR APPLICATION NUMBER:	60/081203
61	PRIOR FILING DATE:	1998-04-09
62	PRIOR APPLICATION NUMBER:	60/081229
63	PRIOR FILING DATE:	1998-04-09
64	PRIOR APPLICATION NUMBER:	60/081955
65	PRIOR FILING DATE:	1998-04-15
66	PRIOR APPLICATION NUMBER:	60/081847
67	PRIOR FILING DATE:	1998-04-15
68	PRIOR APPLICATION NUMBER:	60/081849
69	PRIOR FILING DATE:	1998-04-15
70	PRIOR APPLICATION NUMBER:	60/081952
71	PRIOR FILING DATE:	1998-04-15
72	PRIOR APPLICATION NUMBER:	60/081838
73	PRIOR FILING DATE:	1998-04-15

PRIOR APPLICATION NUMBER: 60/082568	
PRIOR FILING DATE: 1998-04-21	
PRIOR APPLICATION NUMBER: 60/082569	
PRIOR FILING DATE: 1998-04-21	
PRIOR APPLICATION NUMBER: 60/082704	
PRIOR FILING DATE: 1998-04-22	
PRIOR APPLICATION NUMBER: 60/082804	
PRIOR FILING DATE: 1998-04-22	
PRIOR APPLICATION NUMBER: 60/082700	
PRIOR FILING DATE: 1998-04-22	
PRIOR APPLICATION NUMBER: 60/082797	
PRIOR FILING DATE: 1998-04-22	
PRIOR APPLICATION NUMBER: 60/082766	
PRIOR FILING DATE: 1998-04-23	
PRIOR APPLICATION NUMBER: 60/083356	
PRIOR FILING DATE: 1998-04-27	
PRIOR APPLICATION NUMBER: 60/083322	
PRIOR FILING DATE: 1998-04-28	
PRIOR APPLICATION NUMBER: 60/083352	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083455	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083499	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083558	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083559	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083500	
PRIOR FILING DATE: 1998-04-29	
PRIOR APPLICATION NUMBER: 60/083742	
PRIOR FILING DATE: 1998-04-30	
PRIOR APPLICATION NUMBER: 60/084366	
PRIOR FILING DATE: 1998-05-05	
PRIOR APPLICATION NUMBER: 60/084414	
PRIOR FILING DATE: 1998-05-06	
PRIOR APPLICATION NUMBER: 60/084411	
PRIOR FILING DATE: 1998-05-06	
PRIOR APPLICATION NUMBER: 60/084637	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/084639	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/084640	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/084598	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/084600	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/084627	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/084643	
PRIOR FILING DATE: 1998-05-07	
PRIOR APPLICATION NUMBER: 60/085339	
PRIOR FILING DATE: 1998-05-13	
PRIOR APPLICATION NUMBER: 60/085338	
PRIOR FILING DATE: 1998-05-13	
PRIOR APPLICATION NUMBER: 60/085333	
PRIOR FILING DATE: 1998-05-13	
PRIOR APPLICATION NUMBER: 60/085582	
PRIOR FILING DATE: 1998-05-15	
PRIOR APPLICATION NUMBER: 60/085700	
PRIOR FILING DATE: 1998-05-15	
PRIOR APPLICATION NUMBER: 60/085689	
PRIOR FILING DATE: 1998-05-15	
PRIOR APPLICATION NUMBER: 60/085579	
PRIOR FILING DATE: 1998-05-15	
PRIOR APPLICATION NUMBER: 60/085580	

;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697
Query Match 46.3%; Score 50; DB 9; Length 525;
Best Local Similarity 72.7%; Pred. No. 3.2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 10 LSTFMSPCSVT 20
Db 441 LAOEWSPECTVT 451
RESULT 6
US-09-978-697-301
; Sequence 301, Application US/09978697
; Patent No. US20020169284A1
GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Peoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William T.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12

;; PRIOR APPLICATION NUMBER: 60/078004
;; PRIOR FILING DATE: 1998-03-13
;; PRIOR APPLICATION NUMBER: 60/078886
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078936
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081858
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804

PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082700
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082797
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082796
 PRIOR FILING DATE: 1998-04-23
 PRIOR APPLICATION NUMBER: 60/083336
 PRIOR FILING DATE: 1998-04-27
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083392
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083495
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083496
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083499
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083545
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083554
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083558
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083559
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083500
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083742
 PRIOR FILING DATE: 1998-04-30
 PRIOR APPLICATION NUMBER: 60/084366
 PRIOR FILING DATE: 1998-05-05
 PRIOR APPLICATION NUMBER: 60/084414
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084441
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084639
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084640
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084598
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-5-07
 PRIOR APPLICATION NUMBER: 60/084627
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084643
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085339
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085338
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085323
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085582
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085700
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085689
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085580
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085573
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085704
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085697

Query Match 46.3%; Score 50; DB 9; Length 525;
 Best Local Similarity 72.7%; Pred. No. 3.2;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 10. LSTEMSPCSVT 20
 Db 441 LAQEMSPCTVT 451
 RESULT 7
 US-09-978-192A-301
 Sequence 301, Application US/09978192A
 Patent No. US20020177553A1
 GENERAL INFORMATION:
 APPLICANT: Ashkenazi, Avi
 APPLICANT: Baker Kevin P.
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Eaton, Dan
 APPLICANT: Ferrara, Napoleon
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerlitsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin L.
 APPLICANT: Hillan, Kenneth J.
 APPLICANT: Kijavlin, Ivar J.
 APPLICANT: Kuo, Sophia S.
 APPLICANT: Napier, Mary A.
 APPLICANT: Pan, James;
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Shelton, David L.
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 TITLE OF INVENTION: Acids Encoding the Same
 FILE REFERENCE: P2630P1C9
 CURRENT APPLICATION NUMBER: US/09/978,192A
 CURRENT FILING DATE: 2001-10-15
 PRIOR APPLICATION NUMBER: 09/918585
 PRIOR FILING DATE: 2001-07-30
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/064249
 PRIOR FILING DATE: 1997-11-03
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066364
 PRIOR FILING DATE: 1997-11-21
 PRIOR APPLICATION NUMBER: 60/077450
 PRIOR FILING DATE: 1998-03-10
 PRIOR APPLICATION NUMBER: 60/077632
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077641
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077649
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077791
 PRIOR FILING DATE: 1998-03-12
 PRIOR APPLICATION NUMBER: 60/078004
 PRIOR FILING DATE: 1998-03-13
 PRIOR APPLICATION NUMBER: 60/078886
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/078936
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/078910

;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23

;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 46.3%; Score 50; DB 9; Length 525;
Best Local Similarity 72.7%; Pred. No. 3.2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 10 LSTWMSFCSVT 20
DB 441 LAOWMSPTVT 451

```
RESULT 8
US-09-918-171A-21
; Sequence 21, Application US/09918171A
; Patent No. US20020110894A1
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 264/3/04193
; CURRENT APPLICATION NUMBER: US/09/918,171A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/369,364
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-R1
US-09-918-171A-21

Query Match          46.3%; Score 50; DB 10; Length 525;
Best Local Similarity 72.7%; Pred. No. 3.2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 10 LSTWSPCSVT 20
1: |||||:|
Db 441 LAQEMSPCTVT 451

RESULT 9
US-10-052-586-436
; Sequence 436, Application US/10052586
; Patent No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3450R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
```

```
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
```

PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088722
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088863
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876

PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089090
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089908
Query Match 46.3%; Score 50; DB 12; Length 525;
Best Local Similarity 72.7%; Pred. No. 3.2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 10 LSTEMSPCSTVT 20
DB 441 LAQEMSPCTVT 451
RESULT 10
US-09-800-729-89
Sequence 89, Application US/09800729
Patent No. US20020068319A1
GENERAL INFORMATION:
APPLICANT: NI et al.
TITLE OF INVENTION: 32 Human secreted proteins
FILE REFERENCE: P2044P1
CURRENT APPLICATION NUMBER: US/09/800,729
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: PCT/US00/26013
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/155,709
PRIOR FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 217
SOFTWARE: Patentl Ver. 2.0
SEQ ID NO 89
LENGTH: 1745
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-729-89
Query Match 46.3%; Score 50; DB 10; Length 1745;
Best Local Similarity 72.7%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 10 LSTEMSPCSTVT 20
DB 424 LAQEMSPCTVT 434
RESULT 11
US-10-044-807-2
Sequence 2, Application US/10044807
Patent No. US20020165187A1
GENERAL INFORMATION:
APPLICANT: Yu, Xuanchuan
TITLE OF INVENTION: No. US20020165187A1 Human Protease and Polynucleotides Encod
FILE REFERENCE: LEX-0298-USA
CURRENT APPLICATION NUMBER: US/10/044,807
PRIOR FILING DATE: 2002-01-11
PRIOR APPLICATION NUMBER: US 60/261,684
PRIOR FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2

LENGTH: 1762
TYPE: PRT
ORGANISM: homo sapiens
US-10-044-807-2

Query Match
Best Local Similarity 46.3%; Score 50; DB 9; Length 1762;
Best Local Similarity 72.7%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 10 LSTEMSPCSVT 20
1: |||||:11
DB 441 LAQEMSPCTVT 451

RESULT 12
US-09-764-853-554
Sequence 554, Application US/09764853
Patent No. US20020090672A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PJ206
CURRENT APPLICATION NUMBER: US/09/764,853
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 554
LENGTH: 1766
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (333)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-554

Query Match
Best Local Similarity 46.3%; Score 50; DB 10; Length 1766;
Best Local Similarity 72.7%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 10 LSTEMSPCSVT 20
1: |||||:11
DB 445 LAQEMSPCTVT 455

RESULT 13
US-09-862-027-30
Sequence 30, Application US/09862027
Patent No. US20020142428A1
GENERAL INFORMATION:
APPLICANT: Hodges, Martin R.

TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
FILE REFERENCE: 35800/234862
CURRENT APPLICATION NUMBER: US/09/862,027
CURRENT FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 09/345,473
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 82
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 656
TYPE: PRT
ORGANISM: Mus musculus
US-09-862-027-30

Query Match
Best Local Similarity 42.6%; Score 46; DB 10; Length 656;
Best Local Similarity 38.9%; Pred. No. 16;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 EYLKIONSLSTEMSPCS 18
1: |||||:11:11:
DB 570 EHLNPIRENLRGROWKNCA 587

RESULT 14
US-10-042-202-11
Sequence 11, Application US/10042202
Patent No. US20020136733A1
GENERAL INFORMATION:
APPLICANT: Adrian Vivian Sinton HILL, Michael ALDOO,
Catherine Elizabeth Margaret ALLSOPP, Ajit LALVANI, Magdalena
PLEBANSKI, Hilton Carter WHITTLE,
TITLE OF INVENTION: MALARIA PEPTIDES
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: WENDROTH, LIND & PONACK, L.L.P.
STREET: 2033 K Street, N.W., Suite 800,
CITY: Washington
STATE: D.C.,
COUNTRY: U.S.A.
ZIP: 20006-1021

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordpad for Windows 95
CURRENT APPLICATION DATA: US/10/042,202
APPLICATION NUMBER: US/10/042,202
FILING DATE: 11-Jan-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,175
FILING DATE: 28-JAN-1997
APPLICATION NUMBER: NO PCT/GB95/26982
FILING DATE: 30-MAR-1995
APPLICATION NUMBER: GB 9406492.0
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER: 2002_0026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)-721-8200
TELEFAX: (202)-721-8250
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-042-202-11

Query Match
Best Local Similarity 41.7%; Score 45; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 8; 6+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YLKNIONSL 10
1: |||||:11
DB 1 YLKNIONSL 9

RESULT 15
US-09-864-761-40588
Sequence 40588, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeomlca-X-1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:41:41 : Search time 11.8033 Seconds

(without alignments)
162.894 Million cell updates/sec

Title: US-09-931-325C-148

Sequence: 1 EYLKNIQNSLSTWSPSCSVT 20

Scoring table: BLOSUM62
Gapop 10.0 , Capept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	100.0	405	2 S05428	circumsporozoite p
2	99	91.7	424	2 A54533	circumsporozoite p
3	99	91.7	442	2 A54529	circumsporozoite p
4	93	86.1	412	1 02ZQAF	circumsporozoite p
5	91	84.3	388	2 A39756	circumsporozoite p
6	89	82.4	419	1 02ZQAM	circumsporozoite p
7	82	75.9	351	1 02ZQAK	circumsporozoite p
8	82	75.9	363	1 02ZQAK	circumsporozoite p
9	82	75.9	378	1 02ZQAB	circumsporozoite p
10	82	75.9	401	1 02ZQAC	circumsporozoite p
11	80	74.1	378	1 02ZQAL	circumsporozoite p
12	80	74.1	398	1 02ZQAS	circumsporozoite p
13	77	71.3	343	2 A29319	circumsporozoite p
14	77	71.3	367	2 A32068	circumsporozoite p
15	77	71.3	378	1 02ZQAV	circumsporozoite p
16	77	71.3	386	2 A48571	circumsporozoite p
17	77	71.3	387	2 D41156	circumsporozoite p
18	77	71.3	387	2 C41156	circumsporozoite p
19	77	71.3	395	2 A41156	circumsporozoite p
20	77	71.3	429	2 A54504	circumsporozoite p
21	77	71.3	485	2 A60610	circumsporozoite p
22	62	57.4	264	2 A44969	circumsporozoite p
23	60	55.6	332	1 02ZQMB	circumsporozoite p
24	60	55.6	348	1 02ZQMB	circumsporozoite p
25	60	55.6	367	1 02ZQMY	circumsporozoite p
26	52	48.1	388	2 JC6164	circumsporozoite p
27	50	46.3	870	2 A96637	hypothetical prote
28	49	45.4	559	2 S04531	thrombospondin-rel
29	49	45.4	574	2 A46283	sporozoite surface

30	49	45.4	654	2 T32623	hypothetical prote
31	48	44.4	209	2 A41342	circumsporozoite p
32	48	44.4	227	2 B41342	circumsporozoite p
33	48	44.4	1444	2 T18856	angiogenesis inhib
34	46	42.6	156	2 C81868	hypothetical prote
35	46	42.6	469	1 S29126	propeptid precursor
36	46	42.6	551	2 F84567	probable preprotei
37	46	42.6	551	2 A57189	secy protein homol
38	46	42.6	656	2 I49299	receptor interacti
39	46	42.6	1248	2 A47445	reverse gyrase - s
40	45	41.7	38	2 A38869	circumsporozoite p
41	45	41.7	38	2 D60657	circumsporozoite p
42	45	41.7	191	2 JC1524	major allergen m
43	45	41.7	307	2 T45213	superoxide dismuta
44	45	41.7	343	2 T01432	secy protein homol
45	45	41.7	553	2 T01416	secy protein homol

ALIGNMENTS

```
RESULT 1
S05428
circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate NF54)
C:Species: Plasmodium falciparum
C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000
C:Accession: S05428; A45527; I60657
R:Campbell, J.R.
Nucleic Acids Res. 17, 5854, 1989
A:Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate
A:Reference number: S05428; MUID:09345189; PMID:2668895
A:Accession: S05428
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-405 <CAM>
A:Cross-references: EMBL:X15363
R:Capers, P.; Gentz, R.; Matile, H.; Pink, J.R.; Sinigaglia, F.
Mol. Biochem. Parasitol. 35, 185-190, 1989
A:Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate
A:Reference number: A45527; MUID:09345198; PMID:2671723
A:Accession: A45527
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <CAS>
A:Cross-references: GB:M22962; GB:J04650; NID:q160168; PIDN:AAA29527.1; PID:q160169
R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.
Mol. Biochem. Parasitol. 37, 275-280, 1989
A:Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell
A:Reference number: A60657; MUID:90114334; PMID:2481827
A:Accession: I60657
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 319-336, 354-373 <LOC>
A:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:329-383/Domain: thrombospondin type 1 repeat homology <THR>

Query Match 100.0%; Score 108; DB 2; Length 405;
Best local similarity 100.0%; Pred. No. 4; 9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYLKNIQNSLSTWSPSCSVT 20
DB 326 EYLKNIQNSLSTWSPSCSVT 345

RESULT 2
A54533
circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain T4, Thai)
C:Species: Plasmodium falciparum
C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: A54533
R:del Portillo, H.A.; Nussenzweig, R.S.; Enea, V.
Mol. Biochem. Parasitol. 24, 289-294, 1987
```

```

A>Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.
A:Reference number: A54533; MUID:87315205; PMID:3306373
A:Accession: A54533
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <DEL>
A:Cross-references: GB:M19752; NID:g160216; PIDN:AAA29555.1; PID:g160217
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:348-402/domain: thrombospondin type 1 repeat homology <THR1>

Query Match          91.7%; Score 99; DB 2; Length 424;
Best Local Similarity 90.0%; Pred. No. 1.3e+07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 EYLKTIQNSLSTWSPCSVT 20
       :||| ||||||| ||||| |||||
Db      345 QYLKTIQNSLSTWSPCSVT 364

RESULT 3
A54529
circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain Wellcome)
C:Species: Plasmodium falciparum
C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: A54529
R:Lockyer, M.J.; Schwarz, R.T.
Mol. Biochem. Parasitol. 22, 101-108, 1987
A>Title: Strain variation in the circumsporozoite protein gene of Plasmodium falciparum
A:Reference number: A54529; MUID:87115616; PMID:3543671
A:Accession: A54529
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-442 <LOC>
A:Cross-references: GB:M15505; NID:g160214; PIDN:AAA29554.1; PID:g160215
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: tandem repeat
F:366-420/domain: thrombospondin type 1 repeat homology <THR1>

Query Match          91.7%; Score 99; DB 2; Length 442;
Best Local Similarity 90.0%; Pred. No. 1.4e+07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 EYLKTIQNSLSTWSPCSVT 20
       :||| ||||||| ||||| |||||
Db      363 QYLKTIQNSLSTWSPCSVT 382

RESULT 4
A54529
circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate IMT22)
C:Species: Plasmodium falciparum
C>Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jun-2000
C:Accession: A03388
R:Dame, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W.T.
Science 225, 593-599, 1984
A>Title: Structure of the gene encoding the immunodominant surface antigen on the sporozo-
A:Reference number: A03388; MUID:B4250215; PMID:6204383
A:Accession: A03388
A:Molecule type: DNA
A:Residues: 1-412 <DMAM>
A:Cross-references: GB:K02194; NID:g160160; PIDN:AAA29524.1; PID:g160161
A:Experimental source: clone 7G8
C:Comment: Residues 1-16 are the probable signal sequence.
C:Comment: There are 41 copies of a 4-residue repeating unit in the middle domain of the
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:336-390/domain: thrombospondin type 1 repeat homology <THR1>

Query Match          86.1%; Score 93; DB 1; Length 412;
Best Local Similarity 80.0%; Pred. No. 1.1e+06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY      1 EYLKTIQNSLSTWSPCSVT 20
       :||| ||||||| ||||| |||||

```

```

Db      333 QYLKIKNSISTEMSPCSVT 352

RESULT 5
A39756
Circumsporozoite protein - Plasmodium reichenowi
C:Species: Plasmodium reichenowi
C>Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999
C:Accession: A39756
F:Lal, A.A.; Goldman, I.F.
J. Biol. Chem. 266, 6686-6689, 1991
A:Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malar
A:Reference number: A39756; MUID:91201303; PMID:2016283
A:Accession: A39756
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <LAL>
C:Cross-references: GB:M060972; NID:g160228; PIDN:AAA29561.1; PID:g160229
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:312-366/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match      84.3%; Score 91; DB 2; Length 388;
Best Local Similarity 80.0%; Pred. No. 2,1e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 EYLKIKNSISTEMSPCSVT 20
       1 : : : : : : : : : : : : : : : :
Db      309 EFLKIOINNSTEMSPCSVT 328

RESULT 6
OZQAM
Circumsporozoite protein precursor - Plasmodium cynomolgi (strain Mulligan/NH)
N:Alternate names: major sporozoite surface antigen
C:Species: Plasmodium cynomolgi
C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
C:Accession: B26255
F:Galinski, M.R.; Arnott, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Eng
Cell 48, 311-319, 1987
A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
A:Reference number: A90889; MUID:87102878; PMID:3802196
A:Accession: B26255
A:Molecule type: DNA
A:Residues: 1-419 <GAL>
C:Comment: There are three distinct regions in the mature circumsporozoite protein, t
obic membrane-anchoring sequence.
C:Comment: There are 53 tandem copies of a 4-residue repeat.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-419/Product: circumsporozoite protein #status predicted <MAT>
F:99-310/Region: 4-residue repeats
F:344-397/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match      82.4%; Score 89; DB 1; Length 419;
Best Local Similarity 75.0%; Pred. No. 4,7e-06;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 EYLKIKNSISTEMSPCSVT 20
       | | | | | : : : : : | | | | | | |
Db      341 EYLDKIRSTIEMSPCSVT 360

RESULT 7
OZZOKU
Circumsporozoite protein precursor - Plasmodium knowlesi (strain NurI)
N:Alternate names: sporozoite surface protein
C:Species: Plasmodium knowlesi
C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C:Accession: A26253
F:Sharma, S.; Svec, P.; Mitchell, G.H.; Godson, G.N.
Science 229, 779-782, 1985
A:Title: Diversity of circumsporozoite antigen genes from two strains of the malarial

```


A:Reference number: A26253; MUID:85272582; PMID:4023712
A:Accession: A26253
A:Molecule type: DNA
A:Residues: 1-351 <SHA>
A:Cross-references: GB:M1031; NID:g160197; PIDN:AAA29540.1; PID:g160198
C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-351/Product: circumsporozoite protein #status predicted <MAT>
F:98-223/Region: 9-residue repeats
F:224-241/Region: 9-residue repeats
F:276-329/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 75.9%; Score 82; DB 1; Length 351;
Best Local Similarity 65.0%; Pred. No. 4.8e-05;
Matches 13; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Y 1 EYLNKIONSLSTEWSPCSVT 20
Db 273 DYLRKIRSVTTWTPCSVT 292
||:||||:||||:||||
OZZOAK
circumsporozoite protein - Plasmodium knowlesi (strain H)
C:Species: Plasmodium knowlesi
C:Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 16-Jul-1999
A:Accession: A90841; A93315; A03389
R:Galinski, M.R.; Arnott, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Ene, Cell 34, 815-822, 1983
A:Title: Structure of the Plasmodium knowlesi gene coding for the circumsporozoite protein
A:Reference number: A90841; MUID:84026486; PMID:6313209
A:Accession: A90841
A:Molecule type: DNA
A:Residues: 1-363 <OZA>
A:Cross-references: GB:K00822; NID:g160195; PIDN:AAA19699.1; PID:g160196
R:Godson, G.N.; Ellis, J.; Svec, P.; Schlesinger, D.H.; Nussenzweig, V. Nature 305, 29-33, 1983
A:Title: Identification and chemical synthesis of a tandemly repeated immunogenic region
A:Reference number: A93315; MUID:83297689; PMID:6193427
A:Accession: A93315
A:Molecule type: mRNA
A:Residues: 107-208 <GOD>
A:Cross-references: GB:K00772
C:Comment: At least 12 copies of a 12-residue repeating unit occur in this surface protein mosquito to the vertebrate host.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:288-341/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 75.9%; Score 82; DB 1; Length 363;
Best Local Similarity 65.0%; Pred. No. 5e-05;
Matches 13; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Y 1 EYLNKIONSLSTEWSPCSVT 20
Db 285 DYLRKIRSVTTWTPCSVT 304
||:||||:||||:||||
OZZOAB
circumsporozoite protein precursor - Plasmodium cynomolgi (strain Berok)
N:Alternate names: major sporozoite surface antigen
C:Species: Plasmodium cynomolgi
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
A:Accession: D26255
R:Galinski, M.R.; Arnott, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Ene, Cell 48, 311-319, 1987
A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
A:Reference number: A90889; MUID:87102878; PMID:3802196
A:Accession: D26255

A:Molecule type: DNA
A:Residues: 1-378 <GAL>
C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-378/Product: circumsporozoite protein #status predicted <MAT>
F:97-192/Region: 9-residue repeats
F:193-268/Region: 16-residue repeats
F:303-356/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 75.9%; Score 82; DB 1; Length 378;
Best Local Similarity 70.0%; Pred. No. 5.2e-05;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Y 1 EYLNKIONSLSTEWSPCSVT 20
Db 300 EYLRKIRSVTTWTPCSVT 319
||:||||:||||:||||
OZZOAC
circumsporozoite protein precursor - Plasmodium cynomolgi (strain Gombak)
N:Alternate names: major sporozoite surface antigen
C:Species: Plasmodium cynomolgi
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
A:Accession: E26255
R:Galinski, M.R.; Arnott, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Ene, Cell 48, 311-319, 1987
A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
A:Reference number: A90889; MUID:87102878; PMID:3802196
A:Accession: E26255
A:Molecule type: DNA
A:Residues: 1-401 <GAL>
C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-401/Product: circumsporozoite protein #status predicted <MAT>
F:98-278/Region: 11-residue repeats
F:326-379/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 75.9%; Score 82; DB 1; Length 401;
Best Local Similarity 70.0%; Pred. No. 5.6e-05;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Y 1 EYLNKIONSLSTEWSPCSVT 20
Db 323 EYLRKIRSVTTWTPCSVT 342
||:||||:||||:||||
OZZOAL
circumsporozoite protein precursor - Plasmodium cynomolgi (strain London)
N:Alternate names: major sporozoite surface antigen
C:Species: Plasmodium cynomolgi
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
A:Accession: A26255
R:Galinski, M.R.; Arnott, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Ene, Cell 48, 311-319, 1987
A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
A:Reference number: A90889; MUID:87102878; PMID:3802196
A:Accession: A26255
A:Molecule type: DNA
A:Residues: 1-378 <GAL>
C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

Fri Dec 6 14:09:37 2002

us-09-931-325c-148.rpr

Page 5

Search completed: December 6, 2002, 12:46:03
job time : 11.8033 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:38:55 : Search time 6.55738 Seconds

(Without alignments)
126.503 Million cell updates/sec

Title: US-09-931-325C-148

Sequence: 1 EYLKNTQNSLSTFWSPCSVT 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	397	1	CSP_PLAFO
2	99	91.7	424	1	CSP_PLAFT
3	99	91.7	442	1	CSP_PLAFW
4	93	86.1	412	1	CSP_PLAFA
5	91	84.3	388	1	CSP_PLARE
6	89	82.4	419	1	CSP_PLACM
7	82	75.9	351	1	CSP_PLAKU
8	82	75.9	363	1	CSP_PLAKH
9	82	75.9	378	1	CSP_PLAKB
10	80	74.1	398	1	CSP_PLACC
11	79	73.1	401	1	CSP_PLACG
12	77	71.3	343	1	CSP_PLAVS
13	77	71.3	378	1	CSP_PLACL
14	77	71.3	378	1	CSP_PLAVB
15	77	71.3	386	1	CSP_PLASI
16	77	71.3	393	1	CSP_PLABR
17	77	71.3	428	1	CSP_PLAMA
18	60	55.6	339	1	CSP_PLABE
19	60	55.6	347	1	CSP_PLABA
20	60	55.6	367	1	CSP_PLAYO
21	49	45.4	559	1	TRAP_PLAFA
22	48	44.4	1435	1	TR85_HUMAN
23	46	42.6	469	1	PROP_HUMAN
24	46	42.6	551	1	SECY_ARATH
25	46	42.6	656	1	RIK1_MOUSE
26	46	42.6	1248	1	TOPG_STULAC
27	45	41.7	191	1	EXBB_MAIZE
28	45	41.7	247	1	CDH_ENTRL
29	45	41.7	269	1	EXBA_MAIZE
30	45	41.7	307	1	RA31_SCHPO
31	45	41.7	553	1	SECY_MAIZE
32	45	41.7	1074	1	SMSA_HUMAN
33	45	41.7	1077	1	SMSA_MOUSE

34	44	40.7	401	1	TISP_ECOLI
35	44	40.7	450	1	TH12_YEAST
36	44	40.7	849	1	UNCS_ADEEV
37	44	40.7	1173	1	TSP1_XENLA
38	43.5	40.3	2052	1	MY10_BOVIN
39	43	39.8	141	1	V192_FOWPV
40	43	39.8	229	1	NH31_RHORN
41	43	39.8	272	1	VAL1_MSVK
42	43	39.8	272	1	VAL1_MSVN
43	43	39.8	272	1	VAL1_MSVS
44	43	39.8	323	1	VAL1_PASVK
45	43	39.8	626	1	GCKR_RAT

ALIGNMENTS

```

RESULT 1
CSP_PLAFO          STANDARD;          PRT;          397 AA.
ID
AC   P19597; Q25798;
DT   01-FEB-1991 (Rel. 17, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DE   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Circumsporozoite protein precursor (CS).
OS   Plasmodium falciparum (isolate NF54).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5843;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89345189; PubMed=2668895;
RA   Campbell J.R.;
RT   "DNA sequence of the gene encoding a Plasmodium falciparum malaria
RT   candidate vaccine antigen."
RL   Nucleic Acids Res. 17:5854-5854(1989).
RN   [2]
RP   REVISIONS.
RA   Campbell J.R.;
RT   Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RL   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92155296; PubMed=1346766;
RA   Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,
RA   Thomas A.W., Bagar S., Cochran M.A., Thanassi J., Levine M.M.,
RA   Hackett C.S.;
RT   "Plasmodium falciparum: in vitro characterization and human
RT   infectivity of a cloned line."
RL   Exp. Parasitol. 74:159-168(1992).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89364986; PubMed=2671723;
RA   Caspers P., Gentz R., Matile H., Pink J.R., Singagla F.;
RT   "The circumsporozoite protein gene from NF54, a Plasmodium falciparum
RT   isolate used in malaria vaccine trials."
RL   Mol. Biochem. Parasitol. 35:185-190(1989).
CC   -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC   SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
CC   MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC   VERTEBRATE HOST).
CC   -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC   ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC   WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC   -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC   -----
CC   THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
EMBL; X15363; CAA33421.1; -.

```

DR EMBL; M83886; AAA29521.1; -
DR EMBL; M22982; AAA29527.1; -
DR PIR; S05428; S05428.
DR PIR; A45527; A45527.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSPL; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 101 272 43 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT CONFLICT 194 194 A -> ANPNANPNA (IN REF. 4).
SQ SEQUENCE 397 AA; 42646 MW; 9E81146F59EBCA3 CRC64;
Query Match 100.0%; Score 108; DB 1; Length 397;
Best Local Similarity 100.0%; Pred. No. 3,4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 EYLKIONSLSYSTEMSPCSVT 20
318 EYLKIONSLSYSTEMSPCSVT 337
RESULT 2
CSP_PLAFT STANDARD: PRT; 424 AA.
ID CSP_PLAFT P13814;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate 14 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5846;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S., Enea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from Thailand".
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19752; AAA29555.1; -
DR PIR; A54533; A54533.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSPL; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 424 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;

Query Match 91.7%; Score 99; DB 1; Length 424;
Best Local Similarity 90.0%; Pred. No. 8,6e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
DB 1 EYLKIONSLSYSTEMSPCSVT 20
345 QYLKIONSLSYSTEMSPCSVT 364
RESULT 3
CSP_PLAFT STANDARD: PRT; 442 AA.
ID CSP_PLAFT P08307;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87115616; PubMed=3543671;
RA Lockyer M.J., Schwarz R.T.;
RT "Strain variation in the circumsporozoite protein gene of Plasmodium falciparum".
RL Mol. Biochem. Parasitol. 22:101-108(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15505; AAA29554.1; -
DR PIR; A54529; A54529.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSPL; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 442 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 130 320 47 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;
Query Match 91.7%; Score 99; DB 1; Length 442;
Best Local Similarity 90.0%; Pred. No. 8,9e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
DB 1 EYLKIONSLSYSTEMSPCSVT 20
363 QYLKIONSLSYSTEMSPCSVT 382
RESULT 4
CSP_PLAFT STANDARD: PRT; 412 AA.
ID CSP_PLAFT P02893;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN
  SEQUENCE FROM N.A.
RX MEDLINE=64250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;
RT Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum."
RL Science 225:593-599(1984).
CC
  -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
  SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
  MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
  VERTEBRATE HOST).
CC
  -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
  ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
  WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC
  -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; K02194; AAA29524.1; -.
DR PIR; A03388; OZQOAF.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
KM Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.
SQ SEQUENCE 412 AA; 44420 MW; 1EEED3DE90965F8 CRC64;

Query Match 86.1%; Score 93; DB 1; Length 412;
Best Local Similarity 80.0%; Pred. No. 6.9e-07;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYLNKIONSLSTEMSPCSVT 20
DB 333 QYLKIKNSISTEMSPCSVT 352
  :|||:|||||
  |||:|||||

RESULT 5
CSP_PLARE STANDARD: PRT; 388 AA.
ID CSP_PLARE
AC P26694;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91201303; PubMed=2016283;
RA Lal A.A., Goldman I.F.;
RT "Circumsporozoite protein gene from Plasmodium reichenowi, a
RT chimpanzee malaria parasite evolutionarily related to the human
RT malaria parasite Plasmodium falciparum."
RL J. Biol. Chem. 266:6686-6689(1991).
CC
  -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
  SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE

```

```

CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC
  -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
  ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
  WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC
  -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M60972; AAA29561.1; -.
DR PIR; A39756; A39756.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
KM Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 388 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 120 267 62 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 388 AA; 42245 MW; C031EBFBE2335604 CRC64;

Query Match 84.3%; Score 91; DB 1; Length 388;
Best Local Similarity 80.0%; Pred. No. 1.3e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYLNKIONSLSTEMSPCSVT 20
DB 309 EFLKQIONNLSTEMSPCSVT 328
  |||:|||||
  |||:|||||

RESULT 6
CSP_PLACM STANDARD: PRT; 419 AA.
ID CSP_PLACM
AC P08676;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium cynomolgi (strain Mulligan/NIH).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5832;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
RA Nussenzweig R.S., Enea V.;
RT "The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC
  -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
  SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
  MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
  VERTEBRATE HOST).
CC
  -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
  ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
  WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC
  -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M5102; AAA29539.1; -.

```

```
DR PIR: B26255; OZQAM.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSPL; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 419 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 99 314 54 X 4 AA TANDEM REPEATS OF N-A-[DG]-G.
SQ SEQUENCE 419 AA; 38924 MW; 8F46CDD8A1B4EFF4 CRC64;

Query Match 82.4%; Score 89; DB 1; Length 419;
Best Local Similarity 75.0%; Pred. NO. 2.9e-06;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYLKIONSLESTEWPCSVT 20
DB 341 EYLDKIRSTEWPCSVT 360

RESULT 7
CSP_PLAKH STANDARD; PRT; 351 AA.
AC P04922;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium knowlesi (strain H).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid:5852;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-85272582; PubMed-4023712;
RA Sharma S., Svec P., Mitchell G.H., Godson G.N.;
RT "Diversity of circumsporozoite antigen genes from two strains of the
RT malarial parasite Plasmodium knowlesi."
RL Science 229:779-782(1985).
CC -1- SIMILARITY: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M1031; AAA29540.1; -.
DR PIR: A26253; OZQAM.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSPL; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 351 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 101 235 15 X 9 AA TANDEM REPEATS OF A-A-G-A-G-G-
FT E-Q-P.
SQ SEQUENCE 351 AA; 34782 MW; A85E87A152E6485B CRC64;

Query Match 75.9%; Score 82; DB 1; Length 351;
Best Local Similarity 65.0%; Pred. NO. 2.8e-05;
Matches 13; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 EYLKIONSLESTEWPCSVT 20
DB 273 EYLDKIRSTEWPCSVT 292

RESULT 8
CSP_PLAKH STANDARD; PRT; 363 AA.
AC P02894;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium knowlesi (strain H).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid:5851;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-84026486; PubMed-6313209;
RA Ozaki L.S., Svec P., Nussenzweig R.S., Nussenzweig V., Godson G.N.;
RT "Structure of the plasmodium knowlesi gene coding for the
RT circumsporozoite protein."
RL Cell 34:815-822(1983).
RN [2]
RP SEQUENCE OF 84-258 FROM N.A.
RX MEDLINE-83297689; PubMed-6193427;
RA Godson G.N., Ellis J., Svec P., Schlesinger D.H., Nussenzweig V.;
RT "Identification and chemical synthesis of a tandemly repeated
RT immunogenic region of Plasmodium knowlesi circumsporozoite protein."
RL Nature 305:29-33(1983).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K00822; AAA19699.1; -.
DR EMBL: K00772; AAA29556.1; -.
DR PIR: A03389; OZQAM.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSPL; 1.
DR PROSITE: PS50092; TSPL; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 363 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 241 12 X 12 AA TANDEM REPEATS OF N-A-G-Q-P-O-
FT A-Q-G-D-G-A.
SQ SEQUENCE 363 AA; 36793 MW; 574DFA0320A7955 CRC64;

Query Match 75.9%; Score 82; DB 1; Length 363;
Best Local Similarity 65.0%; Pred. NO. 2.9e-05;
Matches 13; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
```



```

RESULT 9
CSP_PLACB
ID CSP_PLACB STANDARD; PRT; 378 AA.
AC P08672;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium cynomolgi (strain Berek).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
RA Nussenzweig R.S., Enea V.;
RT "The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- DOMAIN: THERE ARE 10 TANDEM COPIES OF A 9-RESIDUE REPEAT (PRECEDED BY A 6-RESIDUE INCOMPLETE REPEAT) AND 3 TANDEM COPIES OF A 16-RESIDUE REPEAT (FOLLOWED BY 3 SHORTER, INCOMPLETE COPIES).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M15104; AAA29532.1; -.
DR PIR: D26255; OZ2OAB.
DR InterPro: IPR003067; Crcmsprzoite.
DR Pfam: PF000090; tsp.1; 1.
DR PRINTS: PRO1303; Crcmsprzoite.
DR SMART: SM00209; TSP1.1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 378 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 97 192 10.5 X 9 AA REPEATS.
FT DOMAIN 193 240 3 X 16 AA TANDEM REPEATS.
FT REPEAT 241 251
FT REPEAT 252 260
FT REPEAT 261 268
SQ SEQUENCE 378 AA; 36286 MW; 779BA081C140793F CRC64;

Query Match 75.9%; Score 82; DB 1; Length 378;
Best Local Similarity 70.0%; Pred. No. 3e-05;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYLNKIONSLSTWSPCSVT 20
Db 300 EYLDKIRSTLGEWSPCSVT 319
|||||:|||||

```

```

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5829;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
RA Nussenzweig R.S., Enea V.;
RT "The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M15103; AAA29533.1; -.
DR PIR: C26255; OZ2OAS.
DR InterPro: IPR003067; Crcmsprzoite.
DR Pfam: PF000090; tsp.1; 1.
DR PRINTS: PRO1303; Crcmsprzoite.
DR SMART: SM00209; TSP1.1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 398 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 97 249 17 X 9 AA TANDEM REPEATS OF A-G-N-N-A-A-A-G-E.
SQ SEQUENCE 398 AA; 37718 MW; 6DFA2E8A62ED05BF CRC64;

Query Match 74.1%; Score 80; DB 1; Length 398;
Best Local Similarity 65.0%; Pred. No. 6.5e-05;
Matches 13; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYLNKIONSLSTWSPCSVT 20
Db 320 EYLDKIRSTLGEWSPCSVT 339
|||||:|||||

```

```

RESULT 10
CSP_PLACB
ID CSP_PLACB STANDARD; PRT; 398 AA.
AC P08673;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium cynomolgi (strain Ceylon).

```

CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M15100; AAA29536.1; -.
DR PIR; E26255; OZ20AC.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 401 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 278 17 X 11 AA TANDEM REPEATS OF [DG]-G-A-A-
FT A-A-G-G-G-G-N
FT SEQUENCE 401 AA; 36664 MW; 57D66268238503E CRC64;
Query Match 73.1%; Score 79; DB 1; Length 401;
Best Local Similarity 65.0%; Pred. No. 9.3e-05;
Matches 13; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
OY 1 EYLNKIONSLEWSPSCVT 20
DB 323 EYLDKIRSTIGVEMSPCTVT 342
RESULT 12
CSP_PLAYS
ID CSP_PLAYS STANDARD; PRT; 343 AA.
AC P13826;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein (CS) (Fragment).
OS Plasmodium vivax (strain Salvador I).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=126793;
RN [1]
RP MEDLINE=66070222; PubMed=2416057;
RA McCutchan T.F., Lal A.A., de la Cruz V.F., Miller L.H., Maloy W.L.,
RA Chareonvut V., Beaudoin R.L., Guerry P., Wistar R. Jr., Hoffman S.L.,
RA "Sequence of the immunodominant epitope for the surface protein on
RT sporozoites of Plasmodium vivax.";
RL Science 230:1381-1383(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=87194878; PubMed=2437120;
RA de la Cruz V.F., Lal A.A., Welsh J.A., McCutchan T.F.;
RT "Evolution of the immunodominant domain of the circumsporozoite
RT protein gene from Plasmodium vivax. Implications for vaccines.";
RL J. Biol. Chem. 262:6464-6467(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 1.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 401 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 278 18 X 6 AA TANDEM REPEATS OF D-G-A-R-A-
FT [EA].
FT SEQUENCE 378 AA; 37462 MW; 8295A913636420C5 CRC64;
Query Match 71.3%; Score 77; DB 1; Length 378;
Best Local Similarity 60.0%; Pred. No. 0.00018;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
OY 1 EYLNKIONSLEWSPSCVT 20
DB 300 EYLDKIRSTIGVEMSPCTVT 319

KW Sporozoite; Malaria; Repeat.
FT NON_TER 1 19 X 9 AA TANDEM REPEATS OF P-G-D-R-A-D-
FT DOMAIN 63 243 G-O-P.
FT SEQUENCE 343 AA; 34155 MW; 308EFD5BC15DFC3 CRC64;
Query Match 71.3%; Score 77; DB 1; Length 343;
Best Local Similarity 60.0%; Pred. No. 0.00016;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
OY 1 EYLNKIONSLEWSPSCVT 20
DB 265 EYLDKRVATVGTWTPCSTVT 284
RESULT 13
CSP_PLACL
ID CSP_PLACL STANDARD; PRT; 378 AA.
AC P08675;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium cynomolgi (strain London).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5831;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87102876; PubMed=3802196;
RA Gallinski M.R., Arnot D.E., Cochran A.H., Barnwell J.W.,
RA Nussenzweig R.S., Enea V.;
RA "The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M15101; AAA29537.1; -.
DR PIR; A26255; OZ20AL.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 378 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 211 18 X 6 AA TANDEM REPEATS OF D-G-A-R-A-
FT [EA].
FT SEQUENCE 378 AA; 37462 MW; 8295A913636420C5 CRC64;
Query Match 71.3%; Score 77; DB 1; Length 378;
Best Local Similarity 60.0%; Pred. No. 0.00018;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
OY 1 EYLNKIONSLEWSPSCVT 20
DB 300 EYLDKIRSTIGVEMSPCTVT 319

RESULT 14
CSP_PLAVB STANDARD; PRT; 378 AA.
ID CSP_PLAVB STANDARD; PRT; 378 AA.
AC P08677;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86044510; PubMed=2414847;
RA Annot D.E., Barnwell J.W., Tam J.P., Nussenzweig V., Nussenzweig R.S.,
RT "Circumsporozoite protein of Plasmodium vivax: gene cloning and
RT characterization of the immunodominant epitope.";
RL Science 230:815-818(1985).
RN [2]
RP REVISIONS.
RX MEDLINE=89042133; PubMed=3054880;
RA Annot D.E., Barnwell J.W., Stewart M.J.;
RT "Does biased gene conversion influence polymorphism in the
RT circumsporozoite protein-encoding gene of Plasmodium vivax?";
RL Proc. Natl. Acad. Sci. U.S.A. 85:8102-8106(1988).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M1926; AAA29526.1; -
DR EMBL: J02751; AAA29529.1; ALT_SEQ.
DR PIR: A26256; OZQAV.
DR InterPro: IPR003067; Crcmsprzoite.
DR Pfam: PFR000884; TSP1.
DR Pfam: PFR00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 20 378
FT DOMAIN 97 267
FT 19 X 9 AA TANDEM REPEATS OF P-G-D-R-A-D-
FT G-O-P.
FT CONFLICT 36 36 G -> E (IN REF. 1).
FT CONFLICT 96 96 G -> R (IN REF. 1).
FT CONFLICT 295 295 E -> A (IN REF. 1).
FT CONFLICT 328 328 R -> S (IN REF. 1).
SQ SEQUENCE 378 AA; 37800 MW; C84B5BED053C9ED CRC64;
Query Match 71.3%; Score 77; DB 1; Length 378;
Best Local Similarity 60.0%; Pred. No. 0.00018;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYLNKIONSLSTEMSPCSVT 20
DB 300 EYLDKVRATVGTWTGTEWTPCSVT 319

RESULT 15
CSP_PLAST STANDARD; PRT; 386 AA.
ID CSP_PLAST STANDARD; PRT; 386 AA.
AC Q03110;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium simium.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5859;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93149205; PubMed=8426613;
RA Goldman I.F., Qari S.H., Millet P.G., Collins W.E., Lal A.A.;
RT "Circumsporozoite protein gene of Plasmodium simium, a Plasmodium
RT vivax-like monkey malaria parasite.";
RL Mol. Biochem. Parasitol. 57:177-180(1993).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: I05068; AAA29525.1; -
DR PIR: A48571; A48571.
DR InterPro: IPR003067; Crcmsprzoite.
DR Pfam: PFR000884; TSP1.
DR Pfam: PFR00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
KW Malaria; Sporozoite; Signal; Repeat.
FT SIGNAL 1 19
FT CHAIN 20 386
FT DOMAIN 96 275
FT 20 X 9 AA TANDEM REPEATS OF G-D-R-A-[AD]-
FT G-O-P-A.
SQ SEQUENCE 386 AA; 38567 MW; A0097DADBDE5548DB CRC64;
Query Match 71.3%; Score 77; DB 1; Length 386;
Best Local Similarity 60.0%; Pred. No. 0.00018;
Matches 12; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 EYLNKIONSLSTEMSPCSVT 20
DB 308 EYLDKVRATVGTWTGTEWTPCSVT 327

Search completed: December 6, 2002, 12:44:03
Job time : 7.55738 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 12:41:26 ; Search time 22.9508 Seconds
(without alignments)
179.555 Million cell updates/sec

Title: US-09-931-325c-148
Perfect score: 108
Sequence: 1 EYLNKIONSLSTWSPCSVT 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-rylous:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	108	100.0	80	5	Q9U0P6	Q9U0P6 plasmodium
2	106	98.1	106	5	Q9U0P0	Q9U0P0 plasmodium
3	102	94.4	79	5	Q9U0Q2	Q9U0Q2 plasmodium
4	102	94.4	115	5	Q25835	Q25835 plasmodium
5	102	94.4	436	5	Q27325	Q27325 plasmodium
6	100	92.6	420	5	Q25838	Q25838 plasmodium
7	99	91.7	80	5	Q9U0P8	Q9U0P8 plasmodium
8	99	91.7	80	5	Q9U0P3	Q9U0P3 plasmodium
9	99	91.7	80	5	Q9U0P2	Q9U0P2 plasmodium
10	99	91.7	80	5	Q9U0P0	Q9U0P0 plasmodium
11	99	91.7	80	5	Q9U0Q0	Q9U0Q0 plasmodium
12	99	91.7	80	5	Q9U0P9	Q9U0P9 plasmodium
13	99	91.7	115	5	Q9U034	Q9U034 plasmodium
14	99	91.7	115	5	Q25836	Q25836 plasmodium
15	99	91.7	115	5	Q9U0W7	Q9U0W7 plasmodium
16	99	91.7	117	5	Q25795	Q25795 plasmodium

17	99	91.7	117	5	Q25796	Q25796 plasmodium
18	99	91.7	117	5	Q25797	Q25797 plasmodium
19	99	91.7	408	5	Q25729	Q25729 plasmodium
20	99	91.7	416	5	Q25829	Q25829 plasmodium
21	99	91.7	424	5	Q99256	Q99256 plasmodium
22	99	91.7	424	5	Q27425	Q27425 plasmodium
23	99	91.7	432	5	Q25827	Q25827 plasmodium
24	99	91.7	432	5	Q27246	Q27246 plasmodium
25	99	91.7	442	5	Q25830	Q25830 plasmodium
26	97	89.8	80	5	Q9U0P9	Q9U0P9 plasmodium
27	95	88.0	80	5	Q9U0Q4	Q9U0Q4 plasmodium
28	95	88.0	80	5	Q9U0P1	Q9U0P1 plasmodium
29	95	88.0	80	5	Q9U0P7	Q9U0P7 plasmodium
30	95	88.0	117	5	Q25794	Q25794 plasmodium
31	94	87.0	106	5	Q9U0P2	Q9U0P2 plasmodium
32	93	86.1	80	5	Q9U0Q0	Q9U0Q0 plasmodium
33	93	86.1	80	5	Q9U0P7	Q9U0P7 plasmodium
34	93	86.1	80	5	Q9U0P4	Q9U0P4 plasmodium
35	93	86.1	80	5	Q9U0V9	Q9U0V9 plasmodium
36	92	85.2	73	5	Q25833	Q25833 plasmodium
37	91	84.3	80	5	Q9U0Q1	Q9U0Q1 plasmodium
38	91	84.3	80	5	Q9U0P5	Q9U0P5 plasmodium
39	91	84.3	80	5	Q9U0W1	Q9U0W1 plasmodium
40	91	84.3	115	5	Q25837	Q25837 plasmodium
41	91	84.3	115	5	Q25839	Q25839 plasmodium
42	91	84.3	383	5	Q9GPN1	Q9GPN1 plasmodium
43	91	84.3	420	5	Q25831	Q25831 plasmodium
44	91	84.3	436	5	Q25828	Q25828 plasmodium
45	91	84.3	452	5	Q25834	Q25834 plasmodium

ALIGNMENTS

RESULT 1
ID Q9U0P6 PRELIMINARY; PRT; 80 AA.
AC Q9U0P6:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRAIN-D4264;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains *;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269963; CAB64182.1; -;
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CROMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1 80
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9002 MW; 1CCEAE08E6C9E976 CRC64;
Query Match 100.0%; Score 108; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYLNKIONSLSTWSPCSVT 20
DB 15 EYLNKIONSLSTWSPCSVT 34

RESULT 2

Q9U6P0 PRELIMINARY; PRT; 106 AA.
AC Q9U6P0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1068;
RX MEDLINE=20164888; PubMed=10699256;
RA Allosche A., Silveira H., Conway D.J., Bojang K., Doherty T.,
Cohen J., Pinder M., Greenwood B.M.;
RT "High-throughput sequence typing of T-cell epitope polymorphisms in
Plasmodium falciparum circumsporozoite protein.";
RL Mol. Biochem. Parasitol. 106:273-282(2000).
EMBL: AF181835; AAF03136.1; -;
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PRO1303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1
FT SEQUENCE 106 AA; 11815 MW; 42E13DA19DB0D3C CRC64;

Query Match 98.1%; Score 106; DB 5; Length 106;
Best Local Similarity 95.0%; Pred. No. 1.1e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYLKIONSISTEMSPCSVT 20
Db 43 EYLKIONSISTEMSPCSVT 62

RESULT 3

Q9U0Q2 PRELIMINARY; PRT; 79 AA.
AC Q9U0Q2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DS1;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
EMBL: AJ69945; CAB64243.1; -;
DR InterPro: IPR003067; Circmsprzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PRO1303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1
FT SEQUENCE 79 AA; 8893 MW; 56BFA76D859B416 CRC64;

Query Match 94.4%; Score 102; DB 5; Length 79;
Best Local Similarity 95.0%; Pred. No. 3.5e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYLKIONSISTEMSPCSVT 20
Db 15 EYLKIONSISTEMSPCSVT 34

RESULT 4

Q25835 PRELIMINARY; PRT; 115 AA.
AC Q25835;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=828;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwittes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
EMBL: M83157; AAA29570.1; -;
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PRO1303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1
FT SEQUENCE 115 AA; 12925 MW; 9BE52CFE12BD48F CRC64;

Query Match 94.4%; Score 102; DB 5; Length 115;
Best Local Similarity 95.0%; Pred. No. 5.4e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYLKIONSISTEMSPCSVT 20
Db 36 EYLKIONSISTEMSPCSVT 55

RESULT 5

Q27325 PRELIMINARY; PRT; 436 AA.
AC Q27325;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=837;
RX MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=837;
RA la Cruz V.F.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RC STRAIN=837;

Query Match 94.4%; Score 102; DB 5; Length 115;
Best Local Similarity 95.0%; Pred. No. 5.4e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

RA  Jongwittes S., Tanabe K., Kanbara H.;
RL  Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR  EMBL: M83164; AAA29542.1; -
DR  EMBL: M83150; AAA29563.1; -
DR  EMBL: M83163; AAA29576.1; -
DR  InterPro: IPR003067; Circspprzoite.
DR  InterPro: IPR000884; TSPL.
DR  Pfam: PF00090; tsp_1; 1.
DR  PRINTS: PR01303; CRCMSPRZOITE.
DR  SMART: SM00209; TSPL; 1.
SQ  SEQUENCE 436 AA; 4668 MW; 5B42FF334B68655 CRC64;

Query Match          94.4%; Score 102; DB 5; Length 436;
Best Local Similarity 95.0%; Pred. No. 2.4e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1 EYLNKIONSLSTEWSPCSVT 20
Db  357 EYLNKIONSLSTEWSPCSVT 376

RESULT 6
Q25838      PRELIMINARY;      PRT;      420 AA.
AC  Q25838;
DT  01-NOV-1996 (TREMBlrel. 01, Created)
DT  01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE  01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE  Circumsporozoite protein.
GN  CSP.
OS  Plasmodium falciparum.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5833;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=835B;
RC  MEDLINE=95077069; PubMed=7985759;
RA  Jongwittes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT  "Allelic variation in the circumsporozoite protein of Plasmodium
RT  falciparum from Thai field isolates.";
RL  Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR  EMBL: M83161; AAA29574.1; -
DR  InterPro: IPR003067; Circspprzoite.
DR  InterPro: IPR000884; TSPL.
DR  Pfam: PF00090; tsp_1; 1.
DR  PRINTS: PR01303; CRCMSPRZOITE.
DR  SMART: SM00209; TSPL; 1.
SQ  SEQUENCE 420 AA; 45155 MW; 3A85B92432C2893C CRC64;

Query Match          92.6%; Score 100; DB 5; Length 420;
Best Local Similarity 90.0%; Pred. No. 4.8e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  1 EYLNKIONSLSTEWSPCSVT 20
Db  341 EYLNKIONSLSTEWSPCSVT 360

RESULT 7
Q900P8      PRELIMINARY;      PRT;      80 AA.
AC  Q900P8;
DT  01-MAY-2000 (TREMBlrel. 13, Created)
DT  01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT  01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE  Circumsporozoite protein (Fragment).
GN  CS.
OS  Plasmodium falciparum.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5833;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=MA;


```

```

RA  de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT  "Sequence variation in the non-repeat region of the Plasmodium
RT  falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT  Burmese field isolates and from laboratory strains.";
RL  Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AJ269957; CAB64237.1; -
DR  InterPro: IPR003067; Circspprzoite.
DR  InterPro: IPR000884; TSPL.
DR  Pfam: PF00090; tsp_1; 1.
DR  PRINTS: PR01303; CRCMSPRZOITE.
DR  SMART: SM00209; TSPL; 1.
FT  NON_TER 1
FT  NON_TER 80
SQ  SEQUENCE 80 AA; 9073 MW; ABF40AB8FB142B1E CRC64;

Query Match          91.7%; Score 99; DB 5; Length 80;
Best Local Similarity 90.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  1 EYLNKIONSLSTEWSPCSVT 20
Db  15 EYLNKIONSLSTEWSPCSVT 34

RESULT 8
Q900P3      PRELIMINARY;      PRT;      80 AA.
AC  Q900P3;
DT  01-MAY-2000 (TREMBlrel. 13, Created)
DT  01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT  01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE  Circumsporozoite protein (Fragment).
GN  CS.
OS  Plasmodium falciparum.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5833;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=D4405;
RA  de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT  "Sequence variation in the non-repeat region of the Plasmodium
RT  falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT  Burmese field isolates and from laboratory strains.";
RL  Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AJ269969; CAB64188.1; -
DR  InterPro: IPR003067; Circspprzoite.
DR  InterPro: IPR000884; TSPL.
DR  Pfam: PF00090; tsp_1; 1.
DR  PRINTS: PR01303; CRCMSPRZOITE.
DR  SMART: SM00209; TSPL; 1.
FT  NON_TER 1
FT  NON_TER 80
SQ  SEQUENCE 80 AA; 9033 MW; BA71EBE0DB03193B CRC64;

Query Match          91.7%; Score 99; DB 5; Length 80;
Best Local Similarity 90.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  1 EYLNKIONSLSTEWSPCSVT 20
Db  15 EYLNKIONSLSTEWSPCSVT 34

RESULT 9
Q900P2      PRELIMINARY;      PRT;      80 AA.
AC  Q900P2;
DT  01-MAY-2000 (TREMBlrel. 13, Created)
DT  01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT  01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE  Circumsporozoite protein (Fragment).
GN  CS.
OS  Plasmodium falciparum.


```

```

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5683;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA416;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Thelsen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GURP) from Brazil, Senegalese, and
RT Bunesee field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269970; CAB64189.1;
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSPL; 1.
FT NON_TER 1
FT NON_TER 80
FT NON_TER 80
SQ SEQUENCE 80 AA; 9015 MM; ICF404B8FB142C73 CRC64;
OY
Query Match 91.7%; Score 99; DB 5; Length 80;
Best Local Similarity 90.0%; Pred. No. 1.le-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 EYLKIONSLSTEMSPCSVT 20
15 EYLKIONSLSTEMSPCSVT 34

```

RESULT	10
ID	O9TWO0
PRED	PRELIMINARY;
PT	80 AA.
AC	O9TW00:
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE	Circumsporozoite protein (Fragment).
GN	CS.
OS	Plasmodium falciparum.
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX	NCBI_TaxID=3633;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=FUP/SP,
RA	de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT	"Sequence variation in the non-repeat region of the Plasmodium
RT	falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT	Burmes field isolates and from laboratory strains."
RT	Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR	EMBL; AJ269953; CAB64178.1; -
DR	EMBL; AF269950; CAB64175.1; -
DR	InterPro; IPR003067; Circmsprzoite.
DR	InterPro; IPR000884; TSP1.
DR	Pfam; PF00090; tsp_1; 1.
DR	PRINTS; PRO1303; CRCMSPRZOITE.
DR	SMART; SMO0209; TSP1; 1.
FT	NON_TER
FT	NON_TER
SQ	SEQUENCE 80 AA: 9074 MW: 4614FPDI8FLC0334 CRC64;
Query Match	91.7%; Score 99; DB 5; Length 80;
Best Local Similarity	90.0%; Pred. NO. 1.le+08;
Matches	18; Conservative 1; Mismatches 1; Indels 0; Gaps 0.
OY	1 EYLKIKNSLSTWSPSCVT 20 : iiiiiiiiiii
Db	15 QYLKKIKNSLSTWSPSCVT 34
RESULT	11
ID	O9TV00
PRED	PRELIMINARY;
PT	80 AA.
ID	O9TV00:

[illegible]

RESULT	12	
09TVP9		
ID	09TVP9	
AC	PRELIMINARY;	
DT	01-MAY-2000 (TREMBLrel. 13, Created)	
RT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	
DE	01-JUN-2001 (TREMBLrel. 17, Last annotation update)	
CN	Circumsporozoite protein (Fragment).	
CS	Plasmodium falciparum.	
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
OX	NCBL TaxID=5833;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=M0. M6, M7, AND M1;	
RA	de Stricker K., Vuust J., Jensen S., Oeuvray C., Theisen M.;	
RT	"Sequence variation in the non-repeat region of the Plasmodium falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and Burmese field isolates and from laboratory strains.";	
RL	Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.	
DR	EMBL; AJ269955; CAB64242.1; -	
DR	EMBL; AJ269958; CAB64238.1; -	
DR	EMBL; AJ269959; CAB64239.1; -	
DR	EMBL; AJ269956; CAB64241.1; -	
DR	InterPro: IPR003067; CromeSPRzoite.	
DR	InterPro: IPR000884; TSP1.	
DR	Pfam: PF00090; tsp_1; 1.	
DR	PRINTS: PR01303; CRCMSPRZOITE.	
DR	SMART; SM00209; TSP1; 1.	
FT	NON_TER	
FT	NON_TER	
SQ	SEQUENCE	
Query Match	Best Local Similarity	Score 99; DB 5; Length 80; Pctd. NO. 1.le-08;
Matches	18; Conservative	1; Mismatches 1; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)